

76/14

From: Davis, Natalie
Sent: Thursday, September 19, 2002 3:18 PM
To: STIC-Biotech/ChemLib
Subject: 09/418887

Please search SEQ ID NO: 2 for 09/807470.

Thanks,

Natalie A. Davis, PhD
Patent Examiner
Art Unit 1642
CM1, Rm 8B13
Mailbox 8E12
Ph (703) 308-6410

Edward Hart
Technical Info. Specialist
STIC/Biotech
CMI 6B02 Tel: 305-9203

Pending Nucleic Acid and/or Pending Amino Acid database searches now generate two sets of results. These databases were split into two parts to reduce the time needed to update the databases daily. The split freed up more machine time for processing searches.

Searches run against the Nucleic Acid Pending database produce two sets of results, with the extensions, .rnpm and .rnpn

Searches run against the Amino Acid Pending database produce two sets of results, with the extensions, .rapm and .rapn

The Pending database search results should not be left in the case because they contain data that is confidential.

TYPE OF SEARCH:

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: 9/20/02
Date Completed: 9/20/02
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

NA Sequences: _____
AA Sequences: 1
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)

STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: Q2
WWW/Internet: _____
Other (specify): _____

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2002, 14:34:08 ; Search time 32.48 seconds
(without alignments)
1241.373 Million cell updates/sec

Title: US-09-807-470-2

Perfect score: 1892

Sequence: 1 MGLRKKRNARNPVLSHEFMV.....NGVENPNRIDSPPKKKEKAP 363

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 11107396 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_032802:*
1: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	1892	100.0	363	21	AA198146
2	1892	100.0	363	22	AA198146
3	1463	77.3	369	21	AA198147
4	1463	77.3	369	22	AA198147
5	1463	77.3	369	22	ABG12234
6	1462	77.3	369	22	AA198695
7	1134.5	60.0	416	21	AA1983601
8	922	48.7	304	20	AA198434
9	739	39.1	384	22	AA1985708
10	566	29.9	368	22	ABG65847
11	566	29.9	368	22	ABG66114

12	566	29.9	1575	22	ABG67362	Drosophila melanog
13	414.5	21.9	125	21	AA1980189	Human secreted pro
14	148.5	7.8	411	17	AA1986810	Saccharomyces cere
15	138.5	7.3	394	20	AA1980876	Human LAPH-1 prote
16	138.5	7.3	394	22	AA1980876	Human protein Seq
17	138.5	7.3	394	22	AA1980876	Human protein Seq
18	133.5	7.1	191	21	AA1982489	Human secreted pro
19	117	6.2	488	22	AA1982028	Human polypeptide
20	117	6.2	536	22	AA1980893	Human protein SBO
21	115.5	6.1	296	21	AA1984196	Arabidopsis thalia
22	114.5	6.1	280	21	AA1984197	Arabidopsis thalia
23	111.5	5.9	158	22	AA1981756	Novel signal trans
24	111.5	5.9	181	22	AA19817201	Novel signal trans
25	111	5.9	312	22	AA1982224	Human olfactory re
26	111	5.9	312	22	AA1982549	Human OR-like poly
27	110	5.8	380	20	AA1980877	Human LAPH-2 prote
28	110	5.8	380	22	AA1988560	Human hydrophobic
29	108	5.7	311	22	AA1980782	Human apoptosis pr
30	106.5	5.6	266	21	AA1983806	Human cancer assoc
31	105.5	5.6	310	21	AA19841779	Arabidopsis thalia
32	101.5	5.4	150	21	AA19834938	Human secreted pro
33	98	5.2	232	22	ABG67693	Drosophila melanog
34	97.5	5.2	313	21	AA19833404	zee mays protein f
35	97.5	5.2	397	20	AA19837316	Chlamydia trachoma
36	97.5	5.2	533	22	AA19840417	Human polypeptide
37	96.5	5.1	431	22	ABG30056	Novel human diagno
38	96	5.1	254	21	AA19841780	Arabidopsis thalia
39	93	4.9	313	21	AA1980877	Human G protein-co
40	93	4.9	313	22	AA19824525	Human olfactory re
41	92	4.9	313	22	AA1982073	Human olfactory re
42	91	4.8	307	22	AA1982203	Human olfactory re
43	91	4.8	315	22	AA1982870	Human olfactory re
44	91	4.8	324	22	AA1981539	Human olfactory re
45	90	4.8	309	22	AA19810312	G-protein coupled

ALIGNMENTS

RESULT 1	AA198146
ID	AA198146 standard; Protein; 363 AA.
XX	AC
XX	AA198146;
DT	22-AUG-2000 (first entry)
XX	Rat WAR-1 amino acid sequence.
DE	Endoplasmic reticulum; WAR-1; cancer cell proliferation inhibitor;
XX	diagnosis; cancer; sarcoma; rat.
KW	Rattus norvegicus.
OS	XX
XX	WO200022123-A1.
PN	20-APR-2000.
XX	13-OCT-1999; 99WO-JP05631.
PF	13-OCT-1998; 98JP-0290711.
PR	(SUMU) SUMITOMO PHARM CO LTD.
XX	Tohdoh N, Yoshima T, Komiya K, Tojo S, Nemoto K, Ishikawa H;
XX	Okuyama H;
PI	WPI; 2000-317980/27.
XX	N-PSDB; AAA38012.
DR	Endoplasmic reticulum protein WAR-1 which inhibits cancer cell
XX	proliferation for use in treatment and diagnosis of cancer including
PT	sarcomas of high malignancy

XX
PS
XX
CC
CC
CC
CC
CC
CC
CC
CC
XX
SQ

Claim 1; Fig 2; 89pp; Japanese.

This sequence represents an endoplasmic reticulum protein (WAR-1) amino acid sequence. The invention includes rat and human WAR-1 sequences, expression vectors containing the DNA, cells transformed with the expression vector, antibodies against WAR-1, and probes and primers which hybridize to the DNA encoding WAR-1. The WAR-1 protein inhibits the proliferation of cancer cells, and is used in the treatment and diagnosis of cancers including highly malignant sarcomas.

Sequence 363 AA;

Query Match 100.0%; Score 1892; DB 21; Length 363;
Best Local Similarity 100.0%; Pred. No. 3.8e-213;
Matches 363; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGLRKKARNPPVLSHEPFWQNHADVMSCVGMFFVLGLMFEGETAEMSIIVFTLQHGVTVP 60
DB 1 MGLRKKARNPPVLSHEPFWQNHADVMSCVGMFFVLGLMFEGETAEMSIIVFTLQHGVTVP 60

QY 61 AGLPSGSRITLYHYGVKDLATVFFVLMVAIIHATIQEYVLDKLSRLQLTKGQNKLINE 120
DB 61 AGLPSGSRITLYHYGVKDLATVFFVLMVAIIHATIQEYVLDKLSRLQLTKGQNKLINE 120

QY 121 AQGLSVFYIVSGIMGMIIASENCISDPTLLWKSOPHNMTFQMKFFYISQLAYWFHSFP 180
DB 121 AQGLSVFYIVSGIMGMIIASENCISDPTLLWKSOPHNMTFQMKFFYISQLAYWFHSFP 180

QY 181 ELYFOKVRKQDIPGOLIIYIGLHPLHIGGAYLLYLNLHGLLLMLHYAVELSSVCSLLYF 240
DB 181 ELYFOKVRKQDIPGOLIIYIGLHPLHIGGAYLLYLNLHGLLLMLHYAVELSSVCSLLYF 240

QY 241 GDERYOKGLSLWPVIFISGRVTLIVSVTVGLHAGTNRNGNALSNGVNVLAIAKIAVLS 300
DB 241 GDERYOKGLSLWPVIFISGRVTLIVSVTVGLHAGTNRNGNALSNGVNVLAIAKIAVLS 300

QY 301 SSCSQVYITWTLTWTWQLRWLEDANLHVCGRKRSRSGTENGVENPNRIDSPPKKE 360
DB 301 SSCSQVYITWTLTWTWQLRWLEDANLHVCGRKRSRSGTENGVENPNRIDSPPKKE 360

QY 361 KAP 363
DB 361 KAP 363

RESULT 2
AAB70696
ID AAB70696 standard; Protein; 363 AA.
XX
AC AAB70696;
XX
DT 17-MAY-2001 (first entry)
DE
DE Rat WAR-1 protein sequence SEQ ID NO:2.
XX
KW WAR-1; protein screening; endoplasmic reticulum membrane protein;
KW endoplasmic reticulum membrane transportation; secretory protein;
KW cell membrane protein; cytosolic; CNS active; antiallergic; cancer;
KW antirheumatic; nervous system disorder; immune disorder; allergy;
KW rheumatism; skeletal disorder.
XX
OS Rattus sp.
XX
PN WO200114582-A1.
XX
PD 01-MAR-2001.
XX
PF 17-AUG-2000; 2000WO-JP05488.
XX
PR 20-AUG-1999; 99JP-0234764.
XX

PA (SUMU) SUMITOMO PHARM CO LTD.
XX
PI Tohdoh N, Okuyama H, Imamura M, Ishikawa H, Nemoto K;
XX
DR WPI; 2001-202940/20.
XX
XX N-PSDB; AAF74781.
PT Transformation of a cell with separate vectors expressing the sense and
PT antisense strands of WAR-1 DNA for screening secretory and membrane
PT proteins expressed by the cell -
XX
PS Claim 3; Page 60-62; 79pp; Japanese.
XX
XX The present invention describes a screening method for secretory and
XX membrane proteins consisting of transformation of a cell with separate
XX expression vectors for the sense and antisense RNA of DNA encoding an
XX endoplasmic reticulum membrane protein participating in endoplasmic
XX reticulum transport of proteins. Also described are: (1) secretory and
XX cell membrane proteins identified by the screening method; (2) drug
XX compositions containing these proteins; (3) host cells transformed by
XX the separate expression vectors of the method; and (4) the preparation
XX of secretory and cell membrane proteins by culture of the transformants.
XX The method can be used for the identification and preparation of
XX proteins for use in the treatment and prevention of diseases such as
XX cancer, disorders of the nervous system, immune disorders (including
XX allergies and rheumatism) and skeletal disorders. The present sequence
XX represents a specifically claimed rat WAR-1 protein from the present
XX invention.
SQ Sequence 363 AA;

Query Match 100.0%; Score 1892; DB 22; Length 363;
Best Local Similarity 100.0%; Pred. No. 3.8e-213;
Matches 363; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGLRKKARNPPVLSHEPFWQNHADVMSCVGMFFVLGLMFEGETAEMSIIVFTLQHGVTVP 60
DB 1 MGLRKKARNPPVLSHEPFWQNHADVMSCVGMFFVLGLMFEGETAEMSIIVFTLQHGVTVP 60

QY 61 AGLPSGSRITLYHYGVKDLATVFFVLMVAIIHATIQEYVLDKLSRLQLTKGQNKLINE 120
DB 61 AGLPSGSRITLYHYGVKDLATVFFVLMVAIIHATIQEYVLDKLSRLQLTKGQNKLINE 120

QY 121 AQGLSVFYIVSGIMGMIIASENCISDPTLLWKSOPHNMTFQMKFFYISQLAYWFHSFP 180
DB 121 AQGLSVFYIVSGIMGMIIASENCISDPTLLWKSOPHNMTFQMKFFYISQLAYWFHSFP 180

QY 181 ELYFOKVRKQDIPGOLIIYIGLHPLHIGGAYLLYLNLHGLLLMLHYAVELSSVCSLLYF 240
DB 181 ELYFOKVRKQDIPGOLIIYIGLHPLHIGGAYLLYLNLHGLLLMLHYAVELSSVCSLLYF 240

QY 241 GDERYOKGLSLWPVIFISGRVTLIVSVTVGLHAGTNRNGNALSNGVNVLAIAKIAVLS 300
DB 241 GDERYOKGLSLWPVIFISGRVTLIVSVTVGLHAGTNRNGNALSNGVNVLAIAKIAVLS 300

QY 301 SSCSQVYITWTLTWTWQLRWLEDANLHVCGRKRSRSGTENGVENPNRIDSPPKKE 360
DB 301 SSCSQVYITWTLTWTWQLRWLEDANLHVCGRKRSRSGTENGVENPNRIDSPPKKE 360

QY 361 KAP 363
DB 361 KAP 363

RESULT 3
AAY98147
ID AAY98147 standard; Protein; 369 AA.
XX
AC AAY98147;
XX
DT 22-AUG-2000 (first entry)
XX

DE Human WAR-1 amino acid sequence.
 XX Endoplasmic reticulum; WAR-1; cancer cell proliferation inhibitor;
 KW diagnosis; cancer; sarcoma; human.
 XX
 OS Homo sapiens.
 XX WO200022123-A1.
 XX 20-APR-2000.
 XX 13-OCT-1999; 99WO-JP05631.
 XX 13-OCT-1998; 98JP-0290711.
 XX (SUMI) SUMITOMO PHARM CO LTD.
 XX Tohdoh N, Yoshima T, Komiya K, Tojo S, Nemoto K, Ishikawa H;
 PI Okuyama H;
 XX
 DR WPI; 2000-317980/27.
 DR N-PSDB; AAA38013.
 XX
 PT Endoplasmic reticulum protein WAR-1 which inhibits cancer cell
 PT proliferation for use in treatment and diagnosis of cancer including
 PT sarcomas of high malignancy
 XX
 PS Claim 1; Fig 2; 89pp; Japanese.
 XX
 CC This sequence represents an endoplasmic reticulum protein (WAR-1) amino
 CC acid sequence. The invention includes rat and human WAR-1 sequences,
 CC expression vectors containing the DNA, cells transformed with the
 CC expression vector, antibodies against WAR-1, and probes and primers which
 CC hybridize to the DNA encoding WAR-1. The WAR-1 protein inhibits the
 CC proliferation of cancer cells, and is used in the treatment and diagnosis
 CC of cancers including highly malignant sarcomas.
 XX
 SQ Sequence 369 AA;

Query Match 77.3%; Score 1463; DB 21; Length 369;
 Best Local Similarity 75.8%; Pred. No. 9.6e-163;
 Matches 279; Conservative 43; Mismatches 40; Indels 6; Gaps 5;
 QY 1 MGLRKNARNPPVLSHEFVQNHADMSCVGMFFVLGLMFECTAEMSIYFLTQHGVPVP 60
 DB 1 mglrktknppvlsqefilqnhadivscvgmffilglvfegtaeasivflqlhsvavp 60
 QY 61 -AEGLPSSGRTLYHYGVKDLATVFFYMLVAIIHATIQEYVLDKLSRRLQTLKGQKNKLN 119
 DB 61 aaeeqatgsksllyygvkdlatvffymvliiathiqeyvldklnkrmqtkakqknfn 120
 QY 120 EAGOLSVFVIVSGIMGMIILASENCLSDPTLLKWSOPHNMTFOMKFFYISQLAYWPHSF 179
 DB 121 esqgfsvffvfcvllisencldspdlwkarphsmmtfkmkffysqlaywfhaf 180
 QY 180 PELYFQKRVKODIPGQIYIGLHFGGAYLLYNHGLLMLLHYAVELLSSVCSLLY 239
 DB 181 pelyfaktkkdiprqlvyglnhltgavlylnhlglllvhlyfvellshmcglfy 240
 QY 240 FGDERTQKGLSWPVIIVISGRVLTIVSVVTVGLHAGT-NRNGNALSGNVNVAAKTAV 298
 DB 241 fsdekyqkqislaivalfligrtvltivsvltvgfhlagsqnrndaltgnavvlaaklav 300
 QY 299 LSSSCSIQVYITWLTWVLRWLEDANLHV-CGRKRBSR-SRKGTEN--GVENPNRIDS 354
 DB 301 lssscsiqyvtnltlwlqrwedsnlqascmkksrskkrktengvgvetsnrvc 360
 QY 355 PPKKEKA 362
 DB 361 ppkkeks 368

RESULT 4
 ABG12234
 ID ABG12234 standard; Protein; 369 AA.
 XX
 AC ABG12234;
 XX
 DT 18-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #12225.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX WO200175067-A2.
 XX 11-OCT-2001.
 XX 30-MAR-2001; 2001WO-US08631.
 XX 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI; 2001-639362/73.
 DR N-PSDB; AAS76421.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity
 XX
 PS Claim 20; SEQ ID No 42593; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 369 AA;

Query Match 77.3%; Score 1463; DB 22; Length 369;
 Best Local Similarity 75.8%; Pred. No. 9.6e-163;
 Matches 279; Conservative 43; Mismatches 40; Indels 6; Gaps 5;
 QY 1 MGLRKNARNPPVLSHEFVQNHADMSCVGMFFVLGLMFECTAEMSIYFLTQHGVPVP 60
 DB 1 mglrktknppvlsqefilqnhadivscvgmffilglvfegtaeasivflqlhsvavp 60
 QY 61 -AEGLPSSGRTLYHYGVKDLATVFFYMLVAIIHATIQEYVLDKLSRRLQTLKGQKNKLN 119
 DB 61 aaeeqatgsksllyygvkdlatvffymvliiathiqeyvldklnkrmqtkakqknfn 120

Dd 61 aaeeqatgsksllyygvkdlavtffymvailihatiqeyldklnkrmqftkakqnknf 120

QY 120 EAGOLSVFYVSGIWMGMIILASENCLESDPTLLWKKSOPHNMTFOMKFFYYISQLAYWFHSF 179
:||| ||| ||| ||| ||| ||| :||| :||| ||| ||| ||| ||| :|||

Dd 121 esgfsvfyffsciwgtfllisencisdptllwkarphsmmtfqmkffyisqlaywfha 180

QY 180 PELYTQKYRKODIPCOLIYICGLHFLFHIGGGAYLIYLNLHGILLMLHYAVELLSSVCSELY 239
||||| :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Dd 181 pelytgktkggdiprlqvylghfhigtgayllylahlglllvlyhyveillshmcglfy 240

QY 240 FGDERYQKGLSLWPIVFISGRBLTVLIVSVVTGVLHLAGT--NENGNALSGNVNLAAKIAV 298
||| :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Dd 241 fsdekyqkgislwaivfilgrlvltlvsylvtfvglagsqrnpdalngnrvvlaakiav 300

QY 299 LSSSCSIQYITWLTTTWLQRWLEDANLV-CGRKRGR-SRKTEN--GVENPNRID 354
||||| :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Dd 301 lssstciqaytwnlitlwlrwwedsniqasmckkkrrsskrrtengvgvetsnrvc 360
||||| :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 355 PPKKKKEKA 362

Dd 361 ppkrteks 368

RESULT 5

AAB70695 ID AAB70695 standard; Protein; 369 AA.

XX AC AAB70695;

XX DT 17-MAY-2001 (first entry)

XX DE Human WAR-1 protein sequence SEQ ID NO:1.

KW WAR-1; protein screening; endoplasmic reticulum membrane protein;
KW endoplasmic reticulum membrane transportation; secretory protein;
KW cell membrane protein; cytosolic; CNS active; antiallergic; cancer;
KW anti rheumatic; nervous system disorder; immune disorder; allergy;
KW rheumatism; skeletal disorder.
OS Homo sapiens.
XX WO200114582-A1.
XX PN 01-WAR-2001.
XX PD 17-AUG-2000; 2000WO-JP05488.
XX PF 20-AUG-1999; 99JP-0234764.
XX PR (SUMO) SUMITOMO PHARM CO LTD.
XX PA Tohdoh N, Okuyama H, Imamura M, Ishikawa H, Nemoto K;
PI WPI; 2001-202940/20.
XX DR N-PSDB; AAF74780, RAA74782.
DR DX Transformation of a cell with separate vectors expressing the sense and
PT antisense strands of WAR-1 DNA for screening secretory and membrane
PT proteins expressed by the cell -
PX Claim 2; Page 58-60; 79pp; Japanese.
PS CC The present invention describes a screening method for secretory and
CC membrane proteins consisting of transformation of a cell with separate
CC expression vectors for the sense and antisense RNA of DNA encoding an
CC endoplasmic reticulum membrane protein participating in endoplasmic
CC reticulum transport of proteins. Also described are: (1) secretory and
CC cell membrane proteins identified by the screening method; (2) drug
CC compositions containing these proteins; (3) host cells transformed by
CC the separate expression vectors of the method; and (4) the preparation
CC of secretory and cell membrane proteins by culture of the transformants.
CC The method can be used for the identification and preparation of
CC proteins for use in the treatment and prevention of diseases such as

PT 830 Primers useful for synthesizing full length cDNA clones and their
 XX use in genetic manipulation
 PS Claim 8; SEQ ID NO 2725; 1380pp + sequence listing; English.
 XX
 CC The invention relates to primers for synthesizing full length cDNA
 CC clones. 830 cDNA molecules encoding a human protein have been
 CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
 CC molecules have been determined. Primers for synthesizing the full length
 CC cDNA are useful for clarifying the function of the protein encoded by
 CC the cDNA. The full length clones were obtained by construction of full
 CC length enriched cDNA libraries that were synthesised by the oligo-capping
 CC method. The primers enable the production of the full length cDNA easily
 CC without any special methods. The present sequence is a polypeptide
 CC encoded by a full length human cDNA of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in CD-ROM format directly from EPO.
 XX
 SQ Sequence 369 AA;

Query Match 77.3%; Score 1462; DB 22; Length 369;
 Best Local Similarity 75.5%; Pred. No. 1.3e-162;
 Matches 278; Conservative 44; Mismatches 40; Indels 6; Gaps 5;

QY 1 MGLRKKNARNPPVLSHEFMVQNHADMSCVGMFFVLGLMFEQTAEMSIYVLTQHGVPVP 60
 DB 1 mglrkkstknppvlsqefilghnhdvscvgnffilglvfestaesivfltdhsavp 60
 QY 61 -AGELPSGSTRLYHYGVKDLATVFFVLMVAIIHATIQEYVLDKLSRRLQTKGQNKLN 119
 DB 61 aaeqatgskslvyvgvkdlatvffvmlvaaiihatiqeyvldklnkrmqftkqknfn 120
 QY 120 EAGLSVFIYVSGIMGMITLASENCLSDPTLLWKSOPHNMTFOMKFFVISOLAYWFHSF 179
 DB 121 esqgfvfyfsciwgtfflisencldspdlwarkphsmmfqmkffvysqlaywfhaf 180
 QY 180 PELYFQVRKQDIPGQLIYIGLHFGGAYLLYNLHGLLLMLHYAVELSSVCSLLY 239
 DB 181 pelyfqtktkqdlprqlvyglhfitgalylnhlgllllyhvfvlshmcgify 240
 QY 240 FGERYQKGLSWPIYFISGRVTLIVSVTVGLHAGT-NRNGNALSGVNNVLAAKTAV 298
 DB 241 fdekvyqgslwaifilgrlvltivsvtvgfhlagsqncrpdalgnvnlvlaaktav 300
 QY 299 LSSSCSIQVYITWTLTVMQLRWLEDANLHV-CGRKRRSR-SRKGTEN--GVENPNRIDS 354
 DB 301 lssscsiqayvtwnlltlwlrwvdsnigacsmkkrssrskrtengvgvetsnrvc 360
 QY 355 PPKKKEKA 362
 DB 361 ppkrkeks 368

RESULT 7
 ID AAB43601 standard; Protein; 416 AA.
 AC AAB43601;
 XX
 DT 08-FEB-2001 (first entry)
 XX
 DE Human cancer associated protein sequence SEQ ID NO:1046.
 XX
 KW Human; cancer associated gene; cancer antigen; detection; cancer;
 KW diagnosis; cytostatic; proliferative; vulnerable; immunomodulator;
 KW antidiabetic; antitachymatic; antirheumatic; antithrombotic; antiviral;
 KW antiinflammatory; antihypertensive; antiallergic; antibacterial; cardiant;
 KW dermatological; neuroprotective; thrombolytic; coagulant; nootropic;
 KW vasotropic; antipsoriatic; antidiabetic; gene therapy; inflammation;
 KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
 KW allergic reaction; graft versus host disease; organ rejection;
 KW haemostatic; thrombolytic; cardiovascular disorder; infection;

KW neurological disease; drug screening.
 OS Homo sapiens.
 XX
 PN WO200055350-A1.
 XX
 PD 21-SEP-2000.
 XX
 PF 08-MAR-2000; 2000WO-US05882.
 XX
 PR 12-MAR-1999; 9905-0124270.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM;
 XX
 DR WPI; 2000-587533/55.
 DR N-PSDB; AAC77810.
 XX
 PT Novel isolated nucleic acids comprising sequences encoding peptides
 XX useful for treating or diagnosing e.g. cancer -
 PT
 XX Claim 11; Page 1634-1636; 2352pp; English.
 PS
 XX AAC77607 to AAC78448 encode the human cancer associated proteins given
 CC in AAB43398 to AAB44239. The proteins can have activities based on the
 CC tissues and cells the genes are expressed in. Example of activities
 CC include: cytostatic; proliferative; vulnerable; immunomodulator;
 CC antidiabetic; antitachymatic; antirheumatic; antithrombotic;
 CC antiinflammatory; antihypertensive; antiallergic; antibacterial; antiviral;
 CC dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
 CC nootropic; vasotropic; antipsoriatic and antiangiogenic. The
 CC polynucleotides and polypeptides can be used for preventing, treating or
 CC ameliorating medical conditions and diagnosing pathological conditions.
 CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
 CC the present invention may be used to treat immune disorders by activating
 CC or inhibiting the proliferation, differentiation or mobilisation of
 CC immune cells, to treat disorders of haematopoietic cells, autoimmune
 CC disorders, allergic reactions, graft versus host disease and organ
 CC rejection, modulate haemostatic or thrombolytic activity, modulate
 CC inflammation, cancers, cardiovascular disorders, neurological disease and
 CC bacterial or viral infections. The peptides, nucleotides, antibodies,
 CC agonists and antagonists may be also be used in drug screens. AAC78449 to
 CC AAC78457 and AAB44240 represent sequences used in the exemplification of
 CC the present invention.
 XX
 SQ Sequence 416 AA;

Query Match 60.0%; Score 1134.5; DB 21; Length 416;
 Best Local Similarity 58.7%; Pred. No. 4.5e-124;
 Matches 219; Conservative 61; Mismatches 82; Indels 11; Gaps 4;
 QY 1 MGLRKKNARNPPVLSHEFMVQNHADMSCVGMFFVLGLMFEQTAEMSIYVLTQHGVPVP 60
 DB 43 markkstksppvlshefvlghnhdvscvgnffilglvfestaesivfltdhsavp 102
 QY 61 A-BGLPSGSTRLYHYGVKDLATVFFVLMVAIIHATIQEYVLDKLSRRLQTKGQNKLN 119
 DB 103 ateeqatesvlyyvgvkdlatvffvmlvaaiihatiqeyvldklnkrmqftkqknfn 162
 QY 120 EAGLSVFIYVSGIMGMITLASENCLSDPTLLWKSOPHNMTFOMKFFVISOLAYWFHSF 179
 DB 163 esqgfvfyfsciwgtfflisencldspdlwarkphsmmfqmkffvysqlaywfhaf 222
 QY 180 PELYFQVRKQDIPGQLIYIGLHFGGAYLLYNLHGLLLMLHYAVELSSVCSLLY 239
 DB 223 pelyfqtktkqdlprqlvyglhfitgalylnhlgllllyhvfvlshmcgify 282
 QY 240 FGERYQKGLSWPIYFISGRVTLIVSVTVGLHAGT-NRNGNALSGVNNVLAAKTAV 298
 DB 283 fsnekvyqgslwaifilgrlvltivsvtvgfhlagsqncrpdalgnvnlvlaaktav 342

QY 299 LSSCSIQYIYIWTTLTTLVWLQWLEADANLHVCGRRKRR-----SRSRKGTEGVE---NP 349
 Db 343 lasicvtqfmmwkwfinfglrwrhsafapavkktptvtkg:sskkgtegvngtlts 402
 QY 350 NRIDSPPKKEKA 362
 Db 403 nvadspnrkkks 415

RESULT 8
 AAY48434
 ID AAY48434 standard; Protein; 304 AA.
 AC AAY48434;
 XX
 DT 08-DEC-1999 (first entry)
 XX Human prostate cancer-associated protein 131.
 DE Expressed sequence tag; EST; prostate; tumor; treatment; gene therapy;
 KW cancer; tissue specificity; human.
 XX
 OS Homo sapiens.
 XX
 PN DE19811194-A1.
 XX
 PD 16-SEP-1999.
 XX
 PF 10-MAR-1998; 98DE-10111194.
 XX
 PR 10-MAR-1998; 98DE-10111194.
 XX
 PA (META-) METAGEN GES GENOMFORSCHUNG MBH.
 XX
 XX Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E, Rosenthal A;
 PI WPI; 1999-519629/44.
 DR N-PSDB; AAZ33531.
 XX
 XX New nucleic acid expressed at high level in normal prostatic tissue and
 PT encoded polypeptides, used to treat cancer and screen for therapeutic
 PT agents
 XX
 PS Claim 25; 174; 194pp; German.
 XX
 CC This invention describes novel nucleic acid sequences (A) that are
 CC expressed at high level in normal prostatic tissue. Polypeptides (I)
 CC encoded by (A) are used: (a) for identifying agents for treatment of
 CC prostatic cancer and (b) for therapy of prostate cancer, optionally
 CC where expressed by gene therapy methods. (A) is also used to isolate
 CC full-length genes (for gene therapy) and for recombinant production of
 CC (I), which can be used to raise specific antibodies. (A) are identified
 CC by assembly of ESTs (expressed sequence tags) before these are analyzed
 CC for expression pattern (tissue specificity). This approach eliminates
 CC many of the false results, as regards tissue specificity, associated
 CC with known methods that use single (usually short) ESTs. AAY48304-Y48456
 CC represent peptides encoded by the expressed sequence tags described in
 CC the method of the invention.
 XX
 SQ Sequence 304 AA;

Query Match 48.78; Score 922; DB 20; Length 304;
 Best Local Similarity 58.18; Pred. No. 2.5e-99;
 Matches 176; Conservative 48; Mismatches 69; Indels 10; Gaps 3;

QY 70 TLHYGVKDLATVFYMLVAIIHATIQEVYVLDKLSRRLOLTGKONKLEAGOLSFFVYI 129
 Db 1 sllyygidrlatvfymlyvailhavqeqmldkinrmhfstkskfnesqqlsafyl 60

QY 130 VSGIWMILASENCLSDPTLLWKQSOPNNMNTFOMKFFYISLAYWPHSPPELFOKVRK 189
 Db 61 facvwtgftllisenyisdptllwrayphnltmtfkmkfyisqlaywlhafpelyfqtktk 120

QY 190 QDIPQLIYIGLHFIHGIGAYLLYLHLGLLLMLHYAVELLSSVCSLLYFGDERYOKGL 249
 Db 121 edipqlyyiglyfhiagaylnlnhlglylvlyhvfeflhihrlyfysnekykqgf 180
 QY 250 SLWPIVFIISGRLLVTLIVSVTVGLHLA-GTNRNGNALSGNVNLAAKIAYLSSSCSIQVY 308
 Db 181 slwavflvrltllslvtlvgfglaraenqkldfatgnfnvlavriavlasicvtqaf 240
 QY 309 ITWTLTTLVWLQWLEADANLHVCGRRKRR-----SRSRKGTEGVE---NPNRIDSPPKKK 359
 Db 241 mmwkwfinfglrwrhsafapavkktptvtkg:sskkgtegvngtltsnvadspnrkk 300
 QY 360 EKA 362
 Db 301 eks 303

RESULT 9
 AAM95708
 ID AAM95708 standard; Protein; 384 AA.
 XX
 AC AAM95708;
 XX
 DT 21-NOV-2001 (first entry)
 XX Human reproductive system related antigen SEQ ID NO: 4366.
 DE Human reproductive system related antigen; reproductive system disorder;
 KW cancer; gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN WO200155320-A2.
 XX
 PD 02-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US01339.
 XX
 PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226868.
 PR 22-AUG-2000; 2000US-0227182.

PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241321.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.

PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 06-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX
DR WPI; 2001-465570/50.
DR N-PSDB; AAL01678.
XX
PT Isolated nucleic acid molecule encoding a reproductive system antigen
PT is used in preventing, treating or ameliorating a medical condition
XX
PS Claim 11; SEQ ID NO 4366; 1297pp + Sequence Listing; English.
XX
CC The present invention provides the protein and coding sequences of a
CC number of human reproductive system related antigens. These can be used
CC in the prevention and treatment of reproductive system disorders,
CC including cancer. The present sequence is a protein of the invention.
XX
SQ Sequence 384 AA;

Query Match 39.1%; Score 739; DB 22; Length 384;

Best Local Similarity 41.8%; Pred. No. 1.1e-77;

Matches 156; Conservative 62; Mismatches 135; Indels 20; Gaps 5;

Qy 5 KKNARNPPVLSHEPMQONHADMVSCVGMFFVLGLMPEGTAEMSIIVLTQHGQVWVAEGL 64

Db 18 rrrtksyplfsqetvihnhdgfcvlvcvlgmfxvtaktailfpqynisvpt--- 74

Qy 65 PSGSRTL-YHYGVKDLATVFFYMLVAIIHATIQEYVLDKLSRRLOLTGCKONKLNAGQ 123

Db 75 -adsetvnyhggkdvllifitillhavvqeyildkislrlhlskvkhskfnesgq 133

Qy 124 LSFYIVSGIWMGMIILASENCLSDPTLLMKSQPHNMWTFQMKFFYISOLAYWPHSPPELY 183

Db 134 lvvfhtsviwcfyvvvtegylnprslwedphvhlpfqvkffycqlaywhalpey 193

Qy 184 FQYRKQDIPQOLIYIGLHLFHIGGAYLLNHLGLLLMLHYAVELLSSVCSLLYFGE 243

Db 194 fqvkreeiprlqlyclviahagayllnlsrlgllllllysteflhtxrlfyfde 253

Qy 244 RYQKGLSLWPIVFTSGRLVTIIVSVTVGLHLAG-TNRNGNALSGMNVNLAIAVLSSS 302

Db 254 nneklfsawaaxfgvtrflfilitiavlaigfglarmenqafdkpgkgnftlrcvlvlliv 313

[illegible]

		Query Match	29.9%	Score 566;	DB 22;	Length 368;
		Best Local Similarity	37.0%;	Pred. No. 2.2e-57;		
		Matches 136;	Conservative 64;	Mismatches 130;	Indels 38;	Gaps 9;
QY	4	RKKNARNPVLSHPEWQNQHADMVSCVMFEFVLGLMFEGTAEMSIVFLTLQHGCV--VWPA	61			
		: : : : : : : : : :				
Db	9	rktknknpilshfevignhadiiscvamfvvgllmnestaatafasafisihhnvsgebps	68			
QY	62	EGLPGSRTLHYGVKDLATVFFFWMVAIIHATIQYVLDKLRRRLQLTTKGKQNKLEA	121			
Db	69	reepgpkptylagikdycaiffvtcltlnmahalqefvicksklshkskfklarfnes	128			

```

Qy 122 QQLSVFIVGIVGCMIIASENCLSDPTLLWKSQPHNMATFQKPFYISQLAYWFHSPE 181
Db 129 gqivafyilflsvvgahvllkegyilgqvaglweqfphdpmisflhkfyfvvglayylhmipe 180
Qy 182 LYFOKVR-KODIPGOLYIIGLHLFHGG-----AYLLYLNLHGLLLMLHYAVELLSSV 234
Db 189 lyfqkiktkeeqpkivh-----siggtflvlaytclsfqrlalvlltlhyfsellshv 242
Qy 235 CSLL--YFGDERYOKGLSLWPIVIFISGRIVTLIVSVVTVGLHAGTNRNGNALSGNVNVL 292
Db 243 fqilgvfdreerlaklrvnnnavflirfatsvlgvltlyyiggy-rslalggli--- 298
Qy 293 AAKTAVLSSCSQTVITWTLTIVWLQWLEDANLHVCGKRKRSRSGKTGENGVENPNRI 352
Db 299 -----algyilvifslteqlrakreakk-----eakreaklqltkkpktpk-- 341
Qy 353 DSPPKKKE 360
Db 342 dkvkrkke 349

RESULT 11
ABB66114
ID: ABB66114 standard; Protein; 368 AA.
XX AC ABB66114;
XX AC ABB66114;
Dt 26-MAR-2002 (first entry)
XX Drosophila_melanogaster polypeptide SEQ ID NO 25134.
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX Drosophila melanogaster.
XX WO200171042-A2.
XX 27-SEP-2001.
XX 23-MAR-2001; 2001WO-US09231.
XX 23-MAR-2000; 2000US-191637P.
XX 11-JUL-2000; 2000US-0614150.
XX (PEKE ) PE CORP NY.
XX Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
XX N-PSDB; ABL10217.
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX Disclosure; SEQ ID NO 25134; 21pp + Sequence Listing; English.
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention
CC is useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL161176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL161175) and the encoded proteins
CC (ABBS7737-ABBS7072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from
CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 368 AA.

```

Disclosure: SEO ID NO 25134: 21pp + Sequence Listing: English.

The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from *Protophila*. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-ABB72072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at at.wipo.int/pub/published_pct_sequences.

Sequence 368 AA;

Query Match 29.9%; Score 566; DB 22; Length 368;
 Best Local Similarity 37.0%; Pred. No. 2.2e-57;
 Matches 136; Conservative 64; Mismatches 130; Indels 38; Gaps 9;

QY 4 RKKARNPPVLSHEFMVONHADMVSCVGMFFVLGLMPEGTAEMSIVFLTLQHGCV--VWPA 61
 DB 9 rktknkppllshfviqnhadliiscvnmfvvglmnestaafasafisllhnnvsgeps 68
 QY 62 EGLPSGSRRTLYHYGVKDLATVFFVYMLVAIIHATIQEYVLDKLSRRLQLTGKQNKLEA 121
 DB 69 reqpygkpytiagiakdycaiffytltciimhaiiqefvldkiskkhskfklarfnes 128
 QY 122 GQLSVFVIVSGINGMIILASENCLSDPTLLWKSOPHNMTFQMKFFYISQLAYWFHSFPE 181
 DB 129 gqlvafyllsfvghavllkegylgvqagwefpdpmpsfllhkfvyvqlayylhmlpe 188
 QY 182 LYFQKVR-KODIPGOLYIGLHLPFHIGG-----AYLLYNHLGLLLMLHYAVELLSSV 234
 DB 189 lyfqkiktkeeqpkivh-----sigsftllvlaytlfqrlalvlltlhyfseallshv 242
 QY 235 CSL--YFGDERYOKGLSLMPVIFISGRVTLIVSVVTVGLHLAGTNRNGNALSNGVNVVL 292
 DB 243 fqligvfdreerlaklrnnnavflirfatsvigtlyyggv-rslalggli--- 298
 QY 293 AAKTAVLSSCSIQVYITWTTLTWLQRLWLEDANLHVCGKRKRKRSGKTENGVPNRI 352
 DB 299 -----alqgylvfsfiteqlrakreakk-----eakreakialqtckpaktpk-- 341
 QY 353 DSPPKKKE 360
 DB 342 dkvrkrke 349

RESULT 12
 ABB67362
 ID ABB67362 standard; Protein; 1575 AA.
 XX AC ABB67362;
 XX DT 26-MAR-2002 (first entry)
 XX DE Drosophila melanogaster polypeptide SEQ ID NO 28878.
 XX KW Drosophila; developmental biology; cell signalling; insecticide;
 XX KW pharmaceutical.
 XX OS Drosophila melanogaster.
 XX PN WO200171042-A2.
 XX PD 27-SEP-2001.
 XX PF 23-MAR-2001; 2001WO-US09231.
 XX PR 23-MAR-2000; 2000US-191637P.
 XX PR 11-JUL-2000; 2000US-0614150.
 XX PA (PEKE) PE CORP NY.
 XX PI Venter JC, Adams M, Li PWD, Myers EW;
 XX WIPI; 2001-656860/75.
 XX DR N-PSDB; ABL11465.
 XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX PS Disclosure; SEQ ID NO 28878; 21pp + Sequence Listing; English.
 XX CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is

CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX Sequence 1575 AA;

Query Match 29.9%; Score 566; DB 22; Length 1575;
 Best Local Similarity 37.0%; Pred. No. 2e-56;
 Matches 136; Conservative 64; Mismatches 130; Indels 38; Gaps 9;

QY 4 RKKARNPPVLSHEFMVONHADMVSCVGMFFVLGLMPEGTAEMSIVFLTLQHGCV--VWPA 61
 DB 1216 rktknkppllshfviqnhadliiscvnmfvvglmnestaafasafisllhnnvsgeps 1275
 QY 62 EGLPSGSRRTLYHYGVKDLATVFFVYMLVAIIHATIQEYVLDKLSRRLQLTGKQNKLEA 121
 DB 1276 reqpygkpytiagiakdycaiffytltciimhaiiqefvldkiskkhskfklarfnes 1335
 QY 122 GQLSVFVIVSGINGMIILASENCLSDPTLLWKSOPHNMTFQMKFFYISQLAYWFHSFPE 181
 DB 1336 gqlvafyllsfvghavllkegylgvqagwefpdpmpsfllhkfvyvqlayylhmlpe 1395
 QY 182 LYFQKVR-KODIPGOLYIGLHLPFHIGG-----AYLLYNHLGLLLMLHYAVELLSSV 234
 DB 1396 lyfqkiktkeeqpkivh-----sigsftllvlaytlfqrlalvlltlhyfseallshv 1449
 QY 235 CSL--YFGDERYOKGLSLMPVIFISGRVTLIVSVVTVGLHLAGTNRNGNALSNGVNVVL 292
 DB 1450 fqligvfdreerlaklrnnnavflirfatsvigtlyyggv-rslalggli--- 1505
 QY 293 AAKTAVLSSCSIQVYITWTTLTWLQRLWLEDANLHVCGKRKRKRSGKTENGVPNRI 352
 DB 1506 -----alqgylvfsfiteqlrakreakk-----eakreakialqtckpaktpk-- 1548
 QY 353 DSPPKKKE 360
 DB 1549 dkvrkrke 1556

RESULT 13
 AAG00189
 ID AAG00189 standard; Protein; 125 AA.
 XX AC AAG00189;
 XX DT 06-OCT-2000 (first entry)
 XX DE Human secreted protein, SEQ ID NO: 4270.
 XX KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 XX KW gene therapy; chromosome mapping.
 XX OS Homo sapiens.
 XX PN EP1033401-A2.
 XX PD 06-SEP-2000.
 XX PF 21-FEB-2000; 2000EP-0200610.
 XX PR 26-FEB-1999; 99US-0122487.
 XX PR (GEST) GENSET.
 XX PI Dumas Milne Edwards J, Duclert A, Giordano J;
 XX

DR WPI: 2000-500381/45.
 DR N-PSDB; AAC00195.
 XX
 PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
 XX
 PS Claim 13; SEQ ID 4270; 71pp + CD-ROM; English.
 XX
 CC The present sequence is a polypeptide encoded by one of a large number
 CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
 CC were prepared from total human RNAs or polyA+ RNAs derived from 30
 CC different tissues. EST sequences usually correspond mainly to the 3'
 CC untranslated region (UTR) of the mRNA because they are often obtained
 CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
 CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in
 CC those cases where longer cDNA sequences have been obtained, the full 5'
 CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
 CC ends and can therefore be used to obtain full length cDNAs and genomic
 CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
 CC chromosome mapping procedures. They are used to obtain upstream
 CC regulatory sequences and to design expression and secretion vectors.
 XX
 SQ Sequence 125 AA;

Query Match 21.9%; Score 414.5; DB 21; Length 125;
 Best Local Similarity 64.0%; Pred. No. 2.8e-40;
 Matches 80; Conservative 24; Mismatches 20; Indels 1; Gaps 1;
 QY 1 MGLRKNRNPVLSHEFMVONHADMVSCVGMFFVLGLMFEGETAEMSVIFLTLOHGVVVP 60
 DB 1 mairkstkppvlshfvlghadivscvamlgilmfektakasiifvtlqynvtlp 60
 QY 61 A-EGLPSSGRTLYHYGVKDLATVFFYMLVAIIHATIQEYVLDKLSRELQTKCKQKNLN 119
 DB 61 ateeqatesvslyyygiklatatvffymilvaihaviqeymldknrmxfsktxhskfn 120
 QY 120 EAGQL 124
 DB 121 esgql 125

RESULT 14
 AAR86810
 ID AAR86810 standard; Protein; 411 AA.

XX AAR86810;

XX 28-MAY-1997 (first entry)

DE Saccharomyces cerevisiae LAG1 protein.

XX Saccharomyces cerevisiae; LAG1; life-span limiting domain;
 KW life-span extending domain; stress tolerance; longevity;
 KW recombinant protein production.

OS Saccharomyces cerevisiae.

XX Key Location/Qualifiers

FT Domain 1..190

FT /note= "life-span limiting domain"

FT 194..411

FT /note= "life-span extending domain"

XX W09533834-AL.

XX 14-DEC-1995.

XX 02-JUN-1995; 95WO-US06725.

XX 08-NOV-1994; 94US-0336031.

PR 03-JUN-1994; 94US-0253875.

XX (RESE) RESEARCH CORP TECHNOLOGIES INC.
 XX Jazwinski SM;
 XX WPI: 1996-040238/04.
 DR N-PSDB; AAT07263.
 XX
 PT Eukaryotic LAG1 gene and protein - controls longevity, stress
 PT tolerance and reproductive capacity of eukaryotic cells, for
 PT improved prodn. of recombinant proteins.
 XX
 PS Claim 23; Page 100-102; 154pp; English.
 XX
 CC The Saccharomyces cerevisiae LAG1 gene product is composed of a
 CC life-span limiting domain (see AAR86812) and a life-span extending
 CC domain (see AAR86811). Overexpression of the LAG1 gene in older
 CC cells has a rejuvenating effect, which not only increases
 CC cellular life span, but also reproductive capacity and cellular
 CC tolerance to stress factors such as starvation and low pH. These
 CC cells may be used for the production of recombinant proteins. By
 CC increasing the life span of recombinant cells, the need for
 CC overexpression of recombinant gene products is avoided, and
 CC therefore any subsequent adverse effects on the host cell.
 XX
 SQ Sequence 411 AA;

Query Match 7.8%; Score 148.5; DB 17; Length 411;
 Best Local Similarity 22.0%; Pred. No. 2.9e-08;
 Matches 68; Conservative 46; Mismatches 106; Indels 89; Gaps 11;
 QY 26 MYSCVGMFFVLGLMFEGETAEMSVIFLTLOHGVVVPAGELPSGSRTRYHYGVKDLATVFFY 85
 DB 91 lvcvysayflsgnrttesnplbmfaivsq-----vsgtds-----yakgikdisfvyf 139
 QY 86 MLVAIIIIHATIQEYVLDKLSRRLQTKCKQKNLNEAGOLSVFYI-VSGINGMILL-ASEN 143
 DB 140 mifftflreflmdvvrftvylntvsehrgkrmlegmaifcvgsgpfglymyhsdl 199
 QY 144 CLSDPTLLKWSQPHNMFTFMKFFVISOLAYW-----FHSFPE 181
 DB 200 wlfktkpmrytypvntnplfkfifygaaifaggaqvclqllekprkdykelyfhvlt 259
 QY 182 L-----YFQVKRKODIPQLIYI-----GLHLFH 205
 DB 260 llllwsyyvfhtkm-----glaiytmvdsdfllsktlnylnsvftpfvfglfvf- 312
 QY 206 IGGAYLLYLNHLGLLLMLHYAVELLSSVCSLLYEGDERYKQKGLSLMPITVTSIGRLVTLI 265
 DB 313 -----fwilylrhvnrilrswvltefrhgyvlnfatqgkwcwisl-pivfv-----ll 361
 QY 266 VSVVTVGLH 274
 DB 362 aalqlvnly 370

RESULT 15

AAAY00876

ID AAAY00876 standard; Protein; 394 AA.

XX AAAY00876;

DT 21-MAY-1999 (first entry)

XX Human LAPH-1 protein sequence.

XX Longevity-assurance protein homologue; LAPH-1; human; cell proliferation;
 KW longevity-assurance protein homologue; LAPH-2; signal transduction;
 KW cell cycle regulation; apoptosis; cellular homeostatic pathway; aging;
 KW cancer; inflammation; autoimmune disease; infection;
 KW neurodegenerative disorder.

Search completed: September 20, 2002, 14:34:51
Job time: 43 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2002, 14:34:08 ; Search time 13.07 Seconds
(without alignments)
678.385 Million cell updates/sec

Title: US-09-807-470-2

Perfect score: 1892

Sequence: 1 MGLRKKARPPVLSHEFMV.....NGVENPNRIDSPPKKKEKAP 363

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2.6/ptodata/2/1aa/5A_COMB.pap:*
- 2: /cgn2.6/ptodata/2/1aa/5B_COMB.pap:*
- 3: /cgn2.6/ptodata/2/1aa/6A_COMB.pap:*
- 4: /cgn2.6/ptodata/2/1aa/6B_COMB.pap:*
- 5: /cgn2.6/ptodata/2/1aa/PCTUS_COMB.pap:*
- 6: /cgn2.6/ptodata/2/1aa/backfiles1.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	148.5	7.8	411	2	US-08-336-031-2
2	148.5	7.8	411	2	US-08-902-853-7
3	148.5	7.8	411	5	PCT-US95-06725-2
4	138.5	7.3	394	2	US-08-902-853-1
5	125	6.6	387	2	US-08-902-853-6
6	110	5.8	380	2	US-08-902-853-3
7	89	4.7	649	4	US-08-800-291B-5
8	89	4.7	649	4	US-08-800-291B-6
9	89	4.7	650	4	US-08-800-291B-4
10	87	4.6	1476	4	US-09-256-703-2
11	87	4.6	1479	2	US-08-951-912-4
12	87	4.6	1479	4	US-09-174-077-4
13	87	4.6	1480	1	US-07-637-621-2
14	87	4.6	1480	1	US-08-136-742A-2
15	87	4.6	1480	1	US-08-135-809A-2
16	87	4.6	1480	1	US-08-466-886-17
17	87	4.6	1480	2	US-08-951-912-2
18	87	4.6	1480	2	US-08-951-912-6
19	87	4.6	1480	2	US-08-469-461-2
20	87	4.6	1480	2	US-08-469-461-4
21	87	4.6	1480	2	US-08-691-605-2
22	87	4.6	1480	2	US-08-453-552A-14
23	87	4.6	1480	3	US-07-890-609-2
24	87	4.6	1480	3	US-07-890-609-4
25	87	4.6	1480	3	US-09-248-026-2
26	87	4.6	1480	4	US-08-469-617-17
27	87	4.6	1480	4	US-08-681-838A-2

28	87	4.6	1480	4	US-08-681-838A-3	Sequence 3, Appli
29	87	4.6	1480	4	US-09-174-077-2	Sequence 2, Appli
30	87	4.6	1480	4	US-09-174-077-6	Sequence 6, Appli
31	87	4.6	1480	5	PCT-US93-11667-2	Sequence 2, Appli
32	87	4.6	1480	6	5240846-5	Patent No. 5240846
33	85	4.5	193	2	US-08-336-031-6	Sequence 6, Appli
34	85	4.5	193	5	PCT-US95-06725-6	Sequence 6, Appli
35	82	4.3	2183	3	US-08-746-111-5	Sequence 5, Appli
36	80	4.2	617	1	US-07-879-617A-11	Sequence 11, Appli
37	80	4.2	617	1	US-08-301-722A-3	Sequence 3, Appli
38	80	4.2	617	1	US-08-240-783B-3	Sequence 3, Appli
39	80	4.2	617	1	US-08-753-985-11	Sequence 11, Appli
40	80	4.2	617	3	US-09-084-813-3	Sequence 3, Appli
41	80	4.2	617	5	PCT-US92-09662-3	Sequence 2, Appli
42	79	4.2	326	4	US-08-986-768-2	Sequence 2, Appli
43	79	4.2	602	1	US-08-295-814E-2	Sequence 2, Appli
44	79	4.2	602	1	US-09-343-361-2	Sequence 2, Appli
45	79	4.2	602	5	PCT-US93-01959-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-08-336-031-2
Sequence 2, Application US/08336031
Patent No. 5817782
GENERAL INFORMATION:
APPLICANT: Jazwinski, S. M.
TITLE OF INVENTION: LAG1: A GENE FOR INCREASING THE LONGEVITY OF EUKARYOTES
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: United States
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/336,031
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/253,875
FILING DATE: 03-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9303Z
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO. 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 411 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-336-031-2

Query Match 7.8%; Score 148.5; DB 2; Length 411;
Best Local Similarity 22.0%; Pred. No. 1.1e-08;
Matches 68; Conservative 46; Mismatches 106; Indels 89; Gaps 11;

QY 26 MVSVCNFFVLGMFEGTMSIVFLTQHGVVVPAEGLPGSRTLYHGVKDLATVFFY 85

Db 91 LVCVTSATFLSGNRSTESPLHMFVAISQ-----VDGTD-----YAKGIKDLVSFFV 139
QY 86 MLVAIIIIHATIQEYVLDKLSRLQTKGNKLNAGOLSVFYI-VSGIMGMILL-ASEN 143
Db 140 MIFFTFLEFLMDVIRPFTVNLVNTSEHRQKRMLEQMYAIFYCGVSGPGLYIMYHSDL 199
QY 144 CLSDPTLLKWSOPHNMFTQMKFFVVISQLAYW-----FHSFPE 181
Db 200 WLFKTKPMRTYPTVITNPFLEKIFYLGOAFAWAOQACVVLQLEKPKRDKYKELVFFHIVT 259
QY 182 L-----YFQVKRKODIPGQLIYI-----GLAIYITMDVSDFFLSKLTNLSVFTPFVGLFVF- 312
Db 260 LLLIWSVVFHFTKM-----GLAIYITMDVSDFFLSKLTNLSVFTPFVGLFVF- 312
QY 206 ICGAYLLYLNHLGLLLMLHYAVELLSSVCSLLYFGDERYOKGSLMPTPIVIFISGRVTLI 265
Db 313 -----FWIYLRHVNNIRILWSVLTEFRHGNVYLNFAATQYKCMISL-PIVFFV-----LI 361
QY 266 VSVTVTVGLH 274
Db 362 AALQLVNLV 370

RESULT 2
US-08-902-853-7
; Sequence 7, Application US/08902853
; Patent No. 5945330
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Shah, Purvi
; APPLICANT: Lal, Preeti
; TITLE OF INVENTION: HUMAN LONGEVITY-ASSURANCE PROTEIN HOMOLOGS
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/902,853
; FILING DATE: Herewith
; CLASSIFICATION: ?
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0345 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 411 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: Genbank
; CLONE: 541568
; US-08-902-853-7

Query Match 7.8%; Score 148.5; DB 2; Length 411;
Best Local Similarity 22.0%; Pred. No. 1.1e-08;
Matches 68; Conservative 46; Mismatches 106; Indels 89; Gaps 11;
QY 26 MYSCVGMFEVLGLMFEETAEISYVFLTLQHGVVVPAEGLSPGSRSLYHYGVKDLATVFFY 85
Db 91 LVCVTSATFLSGNRSTESPLHMFVAISQ-----VDGTD-----YAKGIKDLVSFFV 139
QY 86 MLVAIIIIHATIQEYVLDKLSRLQTKGNKLNAGOLSVFYI-VSGIMGMILL-ASEN 143
Db 140 MIFFTFLEFLMDVIRPFTVNLVNTSEHRQKRMLEQMYAIFYCGVSGPGLYIMYHSDL 199
QY 144 CLSDPTLLKWSOPHNMFTQMKFFVVISQLAYW-----FHSFPE 181
Db 200 WLFKTKPMRTYPTVITNPFLEKIFYLGOAFAWAOQACVVLQLEKPKRDKYKELVFFHIVT 259
QY 182 L-----YFQVKRKODIPGQLIYI-----GLAIYITMDVSDFFLSKLTNLSVFTPFVGLFVF- 312
Db 260 LLLIWSVVFHFTKM-----GLAIYITMDVSDFFLSKLTNLSVFTPFVGLFVF- 312
QY 206 ICGAYLLYLNHLGLLLMLHYAVELLSSVCSLLYFGDERYOKGSLMPTPIVIFISGRVTLI 265
Db 313 -----FWIYLRHVNNIRILWSVLTEFRHGNVYLNFAATQYKCMISL-PIVFFV-----LI 361
QY 266 VSVTVTVGLH 274
Db 362 AALQLVNLV 370

RESULT 3
PCT-US95-06725-2
; Sequence 2, Application PC/TUS9506725
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: LAG1: A GENE FOR INCREASING THE
; TITLE OF INVENTION: LONGEVITY OF EUKARYOTES
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/06725
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/253,875 & 08/336,031
; FILING DATE: 03-JUN-1994 & 08-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 93032
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 411 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-06725-2

Query Match 7.8%; Score 148.5; DB 5; Length 411;
Best Local Similarity 22.0%; Pred. No. 1.1e-08;
Matches 68; Conservative 46; Mismatches 106; Indels 89; Gaps 11;
QY 26 MVSVCVGFVGLMREGTAEMSVFLDLQHGUVVPAEGLPSGSRTHYHGVKDLAVFFY 85
DB 91 LVCVTSAYFLSGNTESNPLHMFVAISYQ-----VDGTD-----YAKGKDLSEVFFY 139
QY 86 MLVAIIHATIQEYVLKLSRLQTRKQKQKLNKAGQSVFYI-VSGIWMIL-ASEN 143
DB 140 MIEFTLREFLMDVIRPFTVYLVNTEHQRKMLQOMYAFYCGVSGPGLYIMVHSDL 199
QY 144 CLSDPTLLWKSQPHNMFTQKFFYISQLAYW-----FHSFPE 181
DB 200 WLFKTKMYRTYPVITNPFELFKFYLGQAFAWAQACVVLQLEKPRKDYKELVPHIHT 259
QY 182 L-----YFQVKRKQDIPGQLIYI-----GLAIYITMDVSDFFLSKTLNLYNSVTFPFVGLFVP- 312
DB 260 LLLIWSVVFFTKM-----GLAIYITMDVSDFFLSKTLNLYNSVTFPFVGLFVP- 312
QY 206 IGGAYLYLNHLGMLLMLHAYVELLSVCSLLYFGDERYKQGLSWPIYFISGRVLTLI 265
DB 313 ----FWIYLRHVNRILWSVLTEFRHGNVNLNFAQOQKWNISL-PIVFN-----LI 361
QY 266 VSVVTGVLH 274
DB 362 AALQLVNDY 370

RESULT 4
US-08-902-853-1
; Sequence 1, Application US/08902853
; Patent No. 5945330
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Shah, Purvi
; APPLICANT: Lal, Preeti
; TITLE OF INVENTION: HUMAN LONGEVITY-ASSURANCE PROTEIN HOMOLOGS
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/902,853
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0345 US
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 394 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: LIVERTUT04
CLONE: 2516821
US-08-902-853-1
Query Match 7.3%; Score 138.5; DB 2; Length 394;
Best Local Similarity 24.7%; Pred. No. 1.5e-07;
Matches 44; Conservative 42; Mismatches 81; Indels 11; Gaps 4;
QY 107 RLQITKQKQKLNKAGQSLVFYIVSGIWMILIASENCISDPTLLWKSQPHNMFTQMKF 166
DB 126 RPQUTK-----KFCASWRFLFLYSSFFGSLVSHESWLNWAPVCMWDYRPNQTLKPSLYW 181
QY 167 FYISQLAYWPHSFPELYFQVKRKQDIPGQLIYIIGLHFGIGAYLYLNHLGMLLMLHY 226
DB 182 WYLLLELGYLLSLRLPFD-VKRDKEQYVHFAVILMTFSYANLLRIGSLVLLHLD 240
QY 227 AVELLSVCSLLYFGDERYKQGLS-----LWPIVFISGRVLTIVSVTVGLHLAGTNR 280
DB 241 SSDYLLEACKMVNY--MYYQQVCDALFLFSFVFYTRLVLPFTQILYTTYYSISNR 296
RESULT 5
US-08-902-853-6
; Sequence 6, Application US/08902853
; Patent No. 5945330
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Shah, Purvi
; APPLICANT: Lal, Preeti
; TITLE OF INVENTION: HUMAN LONGEVITY-ASSURANCE PROTEIN HOMOLOGS
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/902,853
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0345 US
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 387 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; ORIGINAL SOURCE:
; HAPLOTYPE: GenBank
; IMMEDIATE SOURCE:
; LIBRARY: 1675382
; US-08-902-853-6

US-08-902-853-3

325	LIRPYLADMTL----	SEVHVMTGGYATIAGSLGAYISFGIDRTSLINAASVMAAPCALAL	381
QY	175	WPHSPPELYFQVRKQDIPGOLYIGLHLPFHGGAYLLYNHLGLLL---	231
Db	382	SKLVYPEVEESKFREE-----	433
QY	232	SSV-----	259
		CSLLYFGDERYQKGLS-----	259
Db	434	AFGLAVLDFTNAAALSGMDWDIOGLSFQILCSYILRPVAFLMG	476

RESULT 9

GENERAL INFORMATION:
APPLICANT: J.D. Young & C.E. Cass
TITLE OF INVENTION: CDNA ENCODING NUCLEOSIDE TRANSPORTER
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/800.291B
FILING DATE: 13-FEB-1997
CLASSIFICATION: 424

CURRENT APPLICATION DATA: US/08/800,291B
 APPLICATION NUMBER: US/08/800,291B
 FILING DATE: 13-FEB-1997
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/499,314
 FILING DATE: 7-JULY-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Haile, Lisa A.
 REGISTRATION NUMBER: 38,347
 REFERENCE/DOCKET NUMBER: 07254/044W01
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 619/678-5070
 TELEFAX: 619/678-5099
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 650 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-800-291B-4

Query Match 4.7%; Score 89; DB 4; Length 650;
Best Local Similarity 21.9%; Pred. No. 0.2;
Matches 62: Conservative 49; Mismatches 102: Indels

[illegible]

FILE REFERENCE: 200116.403C1
CURRENT APPLICATION NUMBER: US/09/174.077
CURRENT FILING DATE: 1998-10-16
EARLIER APPLICATION NUMBER: US 08/951.912
EARLIER FILING DATE: 1997-10-16
NUMBER OF SEQ ID NOS: 6
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 4
LENGTH: 1479
TYPE: PRT
ORGANISM: Homo sapiens
US-09-174-077-4

Query Match 4.6%; Score 87; DB 4; Length 1479;
Best Local Similarity 20.8%; Pred. No. 1.3;
Matches 83; Conservative 63; Mismatches 112; Indels 142; Gaps 24;

QY 24 ADMVSCVGMFFVLGLMFECTAEMSVFVLQHGCVVPAEGLPSGSRITLYHYGV-----KD 78
DB 922 ADTLAMGFRGLPLVHTLITVSKILHHKMLHSVL---QAPMSTLNTLKAGGILNRFSKD 978
QY 79 LA-----TVF-FYMLVAIIH-----ATIOEYVLDK-----LSRRLQLTGKQ 115
DB 979 IAILDPLLPTIFDFIQLLLIVIGAIYVAVLQPIYFVATVPVIVAFIMLRAYFLQTSQ 1038
QY 116 NK-LNPAQOLSVF-YIVSGIMGMIIASENCSDPTLWKSOPHNMFTQMKFFYISOL- 172
DB 1039 LKOLESEGRSPITHLVTSKGLWTLRA-----FGRQPY-----FETLFHKALNLH 1084
QY 173 -AYWFSFPELYFQVKRKODIPGOLYIYGLHFIHGAYLLYL-----NHLGLLLLM-- 223
DB 1085 TANWF-----LYSLTWFMQRIEMIFV---IFFTAVTIFISLTTGEGEGRVGIILTLAM 1136
QY 224 -----LHYAV-----ELLSSVCSLLYFGD-----ERYQKGL----- 249
DB 1137 NIMSTLQWAVNSIDVSLMRSVSRVFKFIDMPTGKPTKSTKPKYKNGOLSKVMIENSH 1196
QY 250 ----SLWPVIFISGRVLT--IVSVTVGLHLAGTNRNGAL----- 285
DB 1197 VKDDIWP-----SGQMTVKDTAKYTEG-----GNAILNISFSISPGORVGLLG 1243
QY 286 ---SGNVNVAIAKIAVLSSCSIQV-YITWTLTVMQLRW 321
DB 1244 RTGSGKSTLLSAFLRLNTEGEIQDGVSW--DSITLQW 1281

RESULT 13
US-07-637-621-2
Sequence 2, Application US/07637621
Patent No. 5407796
GENERAL INFORMATION:
APPLICANT: cutting, gary
APPLICANT: antonarakis, stylianos e
APPLICANT: kazazian jr., haig h
TITLE OF INVENTION: CYSTIC FIBROSIS MUTATION CLUSTER
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner, Birch, McKie and Beckett
STREET: 1001 G Street, N.W.
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20001
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/637,621
FILING DATE: 19910104
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:

NAME: kagan, sarah a
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 1107.030010
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9100
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1480 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: HOMO SAPIENS
US-07-637-621-2

Query Match 4.6%; Score 87; DB 1; Length 1480;
Best Local Similarity 20.8%; Pred. No. 1.3;
Matches 83; Conservative 63; Mismatches 112; Indels 142; Gaps 24;

QY 24 ADMVSCVGMFFVLGLMFECTAEMSVFVLQHGCVVPAEGLPSGSRITLYHYGV-----KD 78
DB 923 ADTLAMGFRGLPLVHTLITVSKILHHKMLHSVL---QAPMSTLNTLKAGGILNRFSKD 979
QY 79 LA-----TVF-FYMLVAIIH-----ATIOEYVLDK-----LSRRLQLTGKQ 115
DB 980 IAILDPLLPTIFDFIQLLLIVIGAIYVAVLQPIYFVATVPVIVAFIMLRAYFLQTSQ 1039
QY 116 NK-LNPAQOLSVF-YIVSGIMGMIIASENCSDPTLWKSOPHNMFTQMKFFYISOL- 172
DB 1040 LKOLESEGRSPITHLVTSKGLWTLRA-----FGRQPY-----FETLFHKALNLH 1085
QY 173 -AYWFSFPELYFQVKRKODIPGOLYIYGLHFIHGAYLLYL-----NHLGLLLLM-- 223
DB 1086 TANWF-----LYSLTWFMQRIEMIFV---IFFTAVTIFISLTTGEGEGRVGIILTLAM 1137
QY 224 -----LHYAV-----ELLSSVCSLLYFGD-----ERYQKGL----- 249
DB 1138 NIMSTLQWAVNSIDVSLMRSVSRVFKFIDMPTGKPTKSTKPKYKNGOLSKVMIENSH 1197
QY 250 ----SLWPVIFISGRVLT--IVSVTVGLHLAGTNRNGAL----- 285
DB 1198 VKDDIWP-----SGQMTVKDTAKYTEG-----GNAILNISFSISPGORVGLLG 1244
QY 286 ---SGNVNVAIAKIAVLSSCSIQV-YITWTLTVMQLRW 321
DB 1245 RTGSGKSTLLSAFLRLNTEGEIQDGVSW--DSITLQW 1282

RESULT 14
US-08-136-742A-2
Sequence 2, Application US/08136742A
Patent No. 5670488
GENERAL INFORMATION:
APPLICANT: Gregory, R.J., Armentano, D., Couture, L.A., Smith,
APPLICANT: A.E.
TITLE OF INVENTION: GENE THERAPY FOR CYSTIC FIBROSIS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: BRUMBAUGH, GRAVES, DONOHUE & RAYMOND
STREET: 30 ROCKEFELLER PLAZA
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/136,742A

COUNTRY: USA
ZIP: 01701
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/135,809A
FILING DATE: 13-OCT-1993
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: LASSEN, ELIZABETH
REGISTRATION NUMBER: 31,845
REFERENCE/DOCKET NUMBER: IG4-9.12
TELECOMMUNICATION INFORMATION:
TELEPHONE: (508) 872-8400
TELEFAX: (508) 872-5415
INFORMATION FOR SEQ. ID NO.: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1480 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-135-809A-2

Query Match	4.6%	Score 87	DB 1	Length 1480		
Best Local Similarity	20.8%	Pred. No. 1.3				
Matches	83	Conservative	63	Mismatches 112; Indels 142; Gaps		
QY	24	ADWVSCVGMFFVLGMEFGTAEMSI	VFTLQHG	VVPAEGLPSGR	SLTYHYGV	-----KD 78
DB	923	ADTLAMGFFRGLPLVHTLTVSKILH	KMLHSVL	---QAPMSTLNTL	KAGGILNR	SKD 979
QY	79	LA-----TVF-FYMLVAIIH	-----ATIQEYVLDK	-----LSRRLQLTKGKQ	115	
DB	980	IAILDLLPLTIFDFIQLLIVIGAIA	VAVLQPYIFATVPVIVAF	TMLRAYFLQTSQ	1039	
QY	116	NK-LNEAQQLSVF-XIVSGWGMIIA	SENCLSOPTLLWK	SOPHMMT	FQMKFFYISOL-	172
DB	1040	LKOLESEGRSFIPTHVLSL	KGLWTFLRA	-----FGROPY	----FETLFFKHALNH	1085
QY	173	-AYWHSFPFELYFORVRQODIPGOL	IYIGLHFIHG	AYLLYL	-----NHLGULLLM--	223
DB	1086	TANWF-----LYLSTLRWFQMI	EFV	---IFFIATVFISIL	TGEGR	VGIIITFLAM 1137
QY	224	-----LHYAV-----ELLSVCS	LLYFGD	-----ERYQKGL	-----	249
DB	1138	NIMSTLQAVNSSIDVDLSMR	SVSRVFKFIDMP	TGKPTKTKPKYK	QKGLSKVM	IENSH 1197
QY	250	-----SLWPVIFISGR	LVTL--IVSVVTVGL	HLAGTNR	GNAL	-----285
DB	1198	VKKDDIWP-----SGQMVT	KDLTAKYTFG	-----GNAILENIS	FSISPGORV	LLG 1244
QY	286	-----SGNVNVLAKTAIVL	SSCSIQV-YITWTL	TTWVLQ	QRW 321	
DB	1245	RTGSGKSTLLSAFLRLNTEGE	IQIDGVSW--DSITLQ	QQ 1282		

Search completed: September 20, 2002, 14:35:11
Job time: 63 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2002, 14:34:08 ; Search time 20.92 Seconds
(without alignments)
1667.324 Million cell updates/sec

Title: US-09-807-470-2

Perfect score: 1892

Sequence: 1 MGLRKKARNPPVLSHFVW.....NGVENPNRIDSPPKKKEKAP 363

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR71.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1150.5	60.8	374	2	S21736	translocating chai
2	1134.5	60.0	374	2	S30034	translocating chai
3	441.5	23.3	371	2	T19417	hypothetical prote
4	434.5	23.0	373	2	T19419	hypothetical prote
5	148.5	7.8	411	2	S46800	LAG1 protein - yea
6	137	7.2	418	2	S30134	hypothetical prote
7	133	7.0	357	2	T40389	longevity-assuranc
8	125	6.6	390	2	T38012	longevity-assuranc
9	116.5	6.2	696	2	B86726	hypothetical prote
10	107.5	5.7	614	2	A69845	Na/H+ antiporter
11	103.5	5.5	393	2	E95954	Na/H+ antiporter (n
12	102.5	5.4	522	2	B83987	ABC transporter (p
13	101	5.3	308	2	H86268	hypothetical prote
14	100.5	5.3	360	2	T27324	hypothetical prote
15	100.5	5.3	397	2	D71467	probable tyrosine
16	99.5	5.3	355	2	D81729	Mtr/TnaB/Tyro perm
17	99	5.2	286	2	AG0403	anaerobic dimethyl
18	98.5	5.2	722	2	G83685	hypothetical prote
19	97.5	5.2	397	2	T00098	hypothetical prote
20	97	5.1	372	2	AH0703	probable membrane
21	96.5	5.1	534	2	T71698	cytochrome-c oxida
22	95.5	5.0	333	2	A10050	probable ABC trans
23	95.5	5.0	411	2	G90154	arsenite transport
24	95	5.0	370	2	H64926	probable membrane
25	95	5.0	370	2	C90928	hypothetical prote
26	95	5.0	370	2	G85776	hypothetical prote
27	95	5.0	509	2	G81929	probable iron-upta
28	94.5	5.0	540	1	I49454	sterol O-acyltrans
29	94.5	5.0	547	2	T27253	hypothetical prote

probable ubiquinol
hypothetical prote
hypothetical prote
cystic fibrosis tr
hypothetical prote
cytochrome-c oxida
probable membrane
NADH dehydrogenase
amino acid transpo
major facilitator
multidrug resistanc
hypothetical prote
hydrogenase-1 oper
hypothetical prote
conserved hypothet
hypothetical prote

ALIGNMENTS

RESULT 1

S21736

translocating chain-associating membrane protein - dog

C:Species: Canis lupus familiaris (dog)

C:Date: 22-Nov-1993 #sequence_revision 01-Dec-1995 #text_change 24-Sep-1999

C:Accession: S21736

R:Goerlich, D.; Hartmann, E.; Prehn, S.; Rapoport, T.A.

Nature 357, 47-52, 1992

A:Title: A protein of the endoplasmic reticulum involved early in polypeptide transloc

A:Reference number: S21736; MUID:92244357

A:Accession: S21736

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-374 <GOE>

A:Cross-references: EMBL:X63678; NID:g941; PIDN:CAA45217.1; PID:g942

C:Superfamily: translocating chain-associating membrane protein

C:Keywords: transmembrane protein

Query Match 60.8%; Score 1150.5; DB 2; Length 374;
Best Local Similarity 59.8%; Pred. No. 1.8e-92;
Matches 223; Conservative 58; Mismatches 81; Indels 11; Gaps 4;

Qy	1	MGLRKKARNPPVLSHFVWQNHADVMSCVGMFFVLGLMFEFTAEMSIVFLTLQHQVWVP	60
Db	1	MAIRKSTKSPVLSHFVWQNHADIVSCVAMVFLGLMFEITAKASIIIVTLQYNVTL	60
Qy	61	A-EGLPSGSRRLHYGVKDLATVFFVYMLVAIIHATIQEYVLDKLSRRLOLTGKONKLN	119
Db	61	ATEBQATESTSLYYGKDLATVFFVYMLVAIIHATIQEYVLDKLNRRMHFSKTSKFN	120
Qy	120	EAGOLSVFVIVSGIWMIIILASENCLSDPTLLMKSPHNMNTFMKFFYISOLAYWFSH	179
Db	121	ESGOLSVFVIVSGIWMIIILASENCLSDPTLLMKSPHNMNTFMKFFYISOLAYWFSH	180
Qy	180	PELVFQVRKQDIPGOLYIIGLHIFHIGGAYLLYLNHLGLLLMLHYAVELLSVCSLLY	239
Db	181	PELVFQVRKQDIPGOLYIIGLHIFHIGGAYLLYLNHLGLLLMLHYAVELLSVCSLLY	240
Qy	240	FGDERYQKGLSLPIVPIISGRVLTIVSVTVGLHLA-GTNRNGNALSGNVNVAIAKIAV	298
Db	241	FSDEKYQKGLSLVAVLVGLVGLRLTLILSVTVGLVGLAENQKLDIFSAGFNVLAVRIAV	300
Qy	299	LSSSCSQVYITVTLTVVLRQWLEDANLHVCGRKR-----SRSRKGTENGVE---NP	349
Db	301	LASITICQAFMMKMFNFQRLRRREHSTFOAPVKKKPTVTKGRSSRKGTENGVTGTVIS	360
Qy	350	NRIDSPPKKKEKA	362
Db	361	NGADSPNRKKEKS	373

Db 185 SYWIHPPEFYQLKRDREIKQSVQALHIAFIATYFFNTRVGLAITLEYITQLIF 244
Qy 233 SVCSLLYFGDERYQKGLS-----LWPIVFISGRGLVTLIVSVWT--VGLHLAGNRNGNAL 285
Db 245 HIARFAHEVG---RKGLSDPAKLFNGSFVLVRLGSIITIAVMTFWYGLROA-ESPVDIS 300
Qy 286 SGNVNVLAAKTAVLSSSCSQTIVTWTTLVQLRWLEDANLHVCGRKRSRKRGTENG 345
Db 301 AGNFTAVIRLVNLLAVVLLQLQLFLLYSFVV-----FHM-GRFRESNAKKEKKKS 348
Qy 346 VENPNRDSPPKKKEK 361
Db 349 AAAAAAV---PKKEKK 361

RESULT 5

S46800
N:AG1 protein - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein YHL003C
C:Species: Saccharomyces cerevisiae
C:Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 23-Mar-2001
C:Accession: S46800; A54012
R:Favellio, T.
A:Submitted to the EMBL Data Library, June 1994
A:Description: The sequence of S. cerevisiae cosmid 9780.
A:Reference number: S46797
A:Accession: S46800
A:Molecule type: DNA
A:Residues: 1-411 <FAV>
A:Cross-references: EMBL:U10555; NID:G500813; PIDN:AAB68429.1; PID:G500820; MIPS:YHL003C
R:D'Amelio, N.P.; Childress, A.M.; Franklin, D.S.; Kale, S.P.; Pinswasdi, C.; Jazwinski, J. Biol. Chem. 269, 15451-15459, 1994
A:Title: Cloning and characterization of LAG1, a longevity-assurance gene in yeast.
A:Reference number: A54012; MUID:94253121
A:Accession: A54012
A:Molecule type: DNA
A:Residues: 1-172, 'IV', 175-219, 'C', 221-300, 'TEISGI', 314, 'EKQE', 315, 'DSNDNPTE', 324, 'A', 324
'A', 381, 'AQGR', 386, 'L', 388, 'NRLARNER' <DAM>
A:Cross-references: GB:U08133
C:Genetics:
A:Gene: SGD:LAG1
A:Cross-references: SGD:S0000995; MIPS:YHL003C
A:Map position: 8L
C:Function:
A:Description: Involved in dextermutation of longevity
C:Superfamily: hypothetical protein YKL008C
C:Keywords: transmembrane protein

Query Match 7.8%; Score 148.5; DB 2; Length 411;
Best Local Similarity 22.0%; Pred. No. 3.6e-05;
Matches 68; Conservative 46; Mismatches 106; Indels 89; Gaps 11;
Qy 26 MYSCVGMFFVLGLMFEGETAEMSVFLTQHGCVVVPAGLPSGRTLYHYGVKDLATVPFY 85
Db 91 LVCVSYAFSLGSGNTESNPLHMFVAISYQ-----VDGTD-----YAKGDKLSFVFY 139
Qy 86 MLVAIIITHATQEVYVLDKLSRRQLTKGKQKLNKNEAGQLSVFYI-VSGIWMGML-ASEN 143
Db 140 MIFTFELREFLMDVVIREFVYLVNTSEHQKRMLEQMAIFYCGVGGPGLYIMVHSD 199
Qy 144 CLSDPTLLWKSQPHNMFTQMKFFYISQLAYW-----FHSFPE 181
Db 200 WLEKTKPMRYTPYITNPFELFKFYLGOAAFAQAQCVLQLKPKDKYKELVPHHVT 259
Qy 182 L-----YFQVRKQDIPGQLIYI-----GLAHLEH 205
Db 260 LLLIWSVVFHTKM-----GLAIYITMDVSDFFLSLTKLNLNSVTFPFVGLFVF- 312
Qy 206 IGGAYLLYLNHLGLLLMLHYAVELSSVCSLLYFGDERYQKGLSLWPIVFISSGLVTLI 265
Db 313 ----FWIYLRHVIRILWLSVLEFRHEGNYVLNFATQYKCHISL-PIVFW-----LI 361

Qy 266 VSVVTVGLH 274
Db 362 AALQLVNLV 370
RESULT 6
S30134
N:hypothetical protein YKL008C - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein YKL156
C:Species: Saccharomyces cerevisiae
C:Date: 18-Jun-1993 #sequence_revision 18-Jun-1993 #text_change 24-Sep-1999
C:Accession: S30134; S37819; S17017
R:Boyer, J.; Pascolo, S.; Richard, G.F.; Dujon, B.
A:Title: Sequence of a 7.8 kb segment on the left arm of yeast chromosome XI reveals mammalian UOG-1 gene.
A:Reference number: S30132; MUID:93255906
A:Accession: S30134
A:Molecule type: DNA
A:Residues: 1-418 <BOY>
A:Cross-references: GB:S59773; NID:G300231; PIDN:AAC60549.1; PID:G300232
R:Boyer, J.; Pascolo, S.; Richard, G.F.; Ghazvini, M.; Colletaux, L.; Thierry, A.; Mori submitted to the Protein Sequence Database, March 1994
A:Reference number: S37813
A:Accession: S37819
A:Molecule type: DNA
A:Residues: 1-418 <BO2>
A:Cross-references: EMBL:Z28008; NID:G485982; PIDN:CAA81843.1; PID:G485983; MIPS:YKL1
R:Amatruda, J.F.; Gattermeier, D.G.; Cooper, J.A.
A:Submitted to the EMBL Data Library, August 1991
A:Description: Yeast capping protein.
A:Reference number: S17016
A:Accession: S17017
A:Molecule type: DNA
A:Residues: 1-149 <AMA>
A:Cross-references: EMBL:X61398; NID:G455515; PIDN:CAA3670.1; PID:G3445
C:Genetics:
A:Map position: 11L
C:Superfamily: hypothetical protein YKL008C
C:Keywords: transmembrane protein

Query Match 7.2%; Score 137; DB 2; Length 418;
Best Local Similarity 23.5%; Pred. No. 0.00037;
Matches 61; Conservative 44; Mismatches 93; Indels 62; Gaps 11;
Qy 67 GSRTLYHYGVKDLATVFFYMLVAIIITHATQEVYVLDKLSRRQLTKGKQKLNKNEAGQLSV 126
Db 121 GDTNAYKGINDLCFVFYIMFFETFLREFLMDVVIREFVIRFAIRLHVTSKHRIKIMEQYAI 180
Qy 127 FYI-VSGIWMGML-ASENCLSDPTLLWKSQPHNMFTQMKFFYISQLAYW----- 175
Db 181 FYTGVSGPFGIYCMYHSDLMWFFNTKAMRYTPDFTNPFELFKFYLGOAAFAQAQCVL 240
Qy 176 -----FHSPEL-----YFQVRKQDIPGQLIYIIGLHFIHGAYLLYLNH 216
Db 241 QLEKPRKDHNLAFHHIVTLLIWSVVFHTKMGLP---IYITMDVSDFLSFSKTLNV 297
Qy 217 L--GLL-----LLMLHYA-VELSSVCS-----LLYFGDERYQKGLSLWPI 254
Db 298 LDSGLAFPSFAIFVAVIYLRHYNILKILMSVLTFQRTGNYVLNFATQYKCHISL-PI 356
Qy 255 VFISGRVTLIVSVTVGLH 274
Db 357 VFW-----LIGALQLVNLV 370

RESULT 7

T40389
longevity assurance protein homolog SPBC3E7.15c - fission yeast (Schizosaccharomyces
C:Species: Schizosaccharomyces pombe

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Jan-2000
 C:Accession: T40389; T40499
 R:Lyne, M.; Wood, V.; Rajandream, M.A.; Barrell, B.G.; Brown, D.; Churcher, C.M.
 submitted to the EMBL Data Library, May 1998
 A:Reference number: Z21924
 A:Accession: T40389
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-357 <LYN>
 A:Cross-references: EMBL:AL023534; PIDN:CAA19018.1; GSPDB:GN00067; SPDB:SPBC3E7.15c
 A:Experimental source: strain 972h-; cosmid c3E7
 R:Gwilliam, R.; Rajandream, M.A.; Barrell, B.G.; Skelton, J.; Churcher, C.M.
 submitted to the EMBL Data Library, September 1998
 A:Reference number: Z21933
 A:Accession: T40499
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-357 <GWT>
 A:Cross-references: EMBL:AL031534; PIDN:CAA20722.1; GSPDB:GN00067; SPDB:SPBC4F6.02c
 A:Experimental source: strain 972h-; cosmid c4F6
 C:Genetics:
 A:Gene: SPBC3E7.15c; SPBC4F6.02c
 A:Map position: 2

Query Match 7.0%; Score 133; DB 2; Length 357;
 Best Local Similarity 20.5%; Pred. No. 0.00068;
 Matches 48; Conservative 45; Mismatches 117; Indels 24; Gaps 7;
 QY 41 EGTAEISVFTL-----QGVVVPAGLPSGRTLYHYGVKDLATVFFYML 87
 DB 29 ERTWIVPLILLTLLGVFVNPNGYIKYIQL-SYPIPGTNPAGYKGRDLIAFLFYAL 87
 QY 88 VALIIHATIQEVLKLSRRQL-TKGQKNLNEAGQLSVFYIVSGIWMIL-ASENCL 145
 DB 88 FTTFREFIMQIIRIGRHNRIRAPKLRFEEQATCTLYFTVMGSLYVMKQTPMWF 147
 QY 146 SDPTLLKWSQPHNMFTQMKFFYISQLAYWPHSPPELYQ--KVRKODIPQQLIYIGLHL 203
 DB 148 FNTDAFWEYEPHYHVGSKAFYIEAAYIQALVLILQLEKPKR-DFRELVVHIIITL 206
 QY 204 FHIGGAYLLNHLGLLLMLHYAVELLSSVCSLLYFGDERYQKGLSLWPIVI 257
 DB 207 LLIGLSYFHTWIGLAVITMDTSDIWLAKSL-----NYNTVIVYPIFVI 255

RESULT 8
 T38012
 longevity-assurance protein 1 - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
 C:Accession: T38012
 R:Churcher, C.M.
 submitted to the EMBL Data Library, September 1997
 A:Reference number: Z21761
 A:Accession: T38012
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-390 <CHU>
 A:Cross-references: EMBL:Z99256; PIDN:CAB16359.1; GSPDB:GN00066; SPDB:SPAC1A6.09c
 A:Experimental source: strain 972h-; cosmid c1A6
 C:Genetics:
 A:Gene: lag1; SPDB:SPAC1A6.09c
 A:Map position: 1

Query Match 6.6%; Score 125; DB 2; Length 390;
 Best Local Similarity 23.9%; Pred. No. 0.0037;
 Matches 57; Conservative 35; Mismatches 100; Indels 46; Gaps 11;
 QY 65 PGSGRTLYHYGVKDLATVFFYMLVAIIHATIQEVLKLSRRQLTKGQK-----KLN 119
 DB 104 PDGS---YKGPDKACDFIWFVITAFRVIYMDYF-----RPFVLNMGVNRKRKVIIFC 156

QY 120 BAGQLSVFYIVSGIWMILASENCLSDPTLLKWSQPHNMFTQMKFFYISQLAYWPHSF 179
 DB 157 EGGYSFFYLLCFWELGLYIVRSSNWSNEEKLFEDYQYMSPLFKAYLYLQLGFWLQOI 216
 QY 180 PELYQKVRKOD-----IPQLIYI--GLHLFHIGGAYLLYLNHLGLLL-----ML 224
 DB 217 IVLLEQ--RRADHWQMFHAHIVTTCALILSYGNFNLRVGNA--ILYIPDLSDYILSGKML 274
 QY 225 HYAVELLSSVCSLLYFGDERYQKGLSLWPIVFIISGRITLVIVSVTV-----GLHL 275
 DB 275 KYLG--FGKICDYLF-----GIFVANSVYSHRYLFSKILRVVVTNAPEIIGGFHL 322
 RESULT 9
 BB6726
 hypothetical protein yica [imported] - Lactococcus lactis subsp. lactis (strain IL140
 C:Species: Lactococcus lactis subsp. lactis
 C>Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
 C:Accession: B86726
 R:Boletín, A.; Wincker, P.; Mauger, S.; Jallón, O.; Malarre, K.; Weissenbach, J.; Eh
 Genome Res. 11, 731-753, 2001
 A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis
 A:Reference number: A86625; MUID:21235186; PMID:11337471
 A:Accession: B86726
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-696 <STO>
 A:Cross-references: GB:AE005176; PID:gl2727372; PIDN:AAK04908.1; GSPDB:GN00146
 A:Experimental source: strain IL1403
 C:Genetics:
 A:Gene: yica

Query Match 6.2%; Score 116.5; DB 2; Length 696;
 Best Local Similarity 21.9%; Pred. No. 0.041;
 Matches 80; Conservative 62; Mismatches 122; Indels 101; Gaps 18;
 QY 27 VSCVGMFVLGL-MFEGTAEMSIIVFLTLQHVVPABGLPSGRTLYHYGVKDLATVFFY 85
 DB 21 ILICINFILNRFSGYTGDDFLY----HFVYTG--WPEHLREYHNLWDWLAVHHT 74
 QY 86 MLVAIIHATIQEVLKLSRRQLTKGQKNLNEAGQLSVFYIVSGIWMILASENCL 145
 DB 75 ML---IWNARMTSIIFEFA--MQIPKFLNIIN-----SLIYVLIGLLINLVSGKAF 124
 QY 146 SDPTLLKWSQPHNMFTQMKFFYISQLA-----YWFHSFPELYQKVRKODIPG 194
 DB 125 LKPS-----HLSLTFLLMWFFLPFGMSIVLVWSGATNLYLWFLSLVILFLAFREDIAA 177
 QY 195 QLIVYIGLHLF-----HIGG--AYLLYL-----NH-----LGLLLMLH 225
 DB 178 RSNWISLGLFILGLTLGLTNEVGATAFLLALLFTIFNRRQPSERVLTQIFGVLAGIG 237
 QY 226 YAVELLSSV-----CSLLYFGDERYQKGLSLWPIVFIISGRITLVIVS- 267
 DB 238 FFIQLLLSSGSETQNTYKSGAGFLQHLSDVFTGMQY--SGFLLPIILLGLGLYLRIQW 296
 QY 268 -----VVTVGLHAGTNRNGNALSGNVVLAAKI--AVLSSCSIQYVITWTLFTVWL 318
 DB 297 TEKVTLVITSLEFL-----GSALAGSIALASPISPAFLWFPANILLITITLLLL--I 347
 QY 319 QRWLE 323
 DB 348 EAWQE 352
 RESULT 10
 A69845
 Na+/H+ antiporter homolog yjbQ - Bacillus subtilis
 C:Species: Bacillus subtilis
 C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 15-Oct-1999
 C:Accession: A69845

R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berteol
C.; Bron, S.; Bruchli, C.V.; Bruchli, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galled
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon,
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, Y.; Sekowska, A.; Seron
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Winat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
A:Reference number: A69580; MUID:98044033
A:Accession: A69845
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-614 <KUN>
A:Cross-references: GB:299110; GB:AL009126; NID:2633472; PIDN:CAB13021.1; PID:ell83184;
A:Experimental source: strain 168
C:Genetics:
A:Gene: yjbQ

Query Match 5.7%; Score 107.5; DB 2; Length 614;
Best Local Similarity 20.3%; Pred. No. 0.21;
Matches 67; Conservative 58; Mismatches 123; Indels 79; Gaps 14;

Qy 24 ADMVSCVGMF-----EVLGMPECTAEMISVLTQLQGVVPA---EGLPS---GS 68
Db 103 AASVIFGVIFILSLLSYGVLAGFONAFMTLIISTLSLGVVPTLKEERIMNSNGQ 162
Qy 69 RLYHYGVKDLATVFF-----YMLVAIIHATIOEYVLDKLSRRLQLTGK 114
Db 163 IILLVAVIADLATMILLAVFSSLYGDSGNMMLLILFAAGVVLFFGVFRHRSFVQSM 222
Qy 115 QNKLEAGQLSVFIVSGINGMI-ILASENCLS---DPTLLMKSQPHNMTFMKFFEYS 170
Db 223 SKGTIOGTTRAITLIIVIVALSSEGAENILGAFAGVLVSLSPKELVQOLDSF--- 279
Qy 171 QLAWYHFFPELYFQKVRKODIPQOLIYIGLHLFHIGGAYLLYLNHLLMLHLYAVEL 230
Db 280 --GYGFL-----IPFFVMGVKL-----DIWTLFQDKTILIMIPLLLLALL 319
Qy 231 LSSVCSLLYFGDERYQKGLSLW---PIVFISGRLYT---LIVSVTVG--LHLAGTNRN 281
Db 320 VSKTIIPWY-----LKKWYDNRITFASGFLTSTLSLVIAAATIGQQLHVISTN-- 368
Qy 282 GNALSGNVNVLAAKIAVLSLSSCSQIYV 308
Db 369 --MSGALILVAVIASITFTPCFKLY 392

RESULT 11
E96954
Na/H antiporter (napA) [imported] - Clostridium acetobutylicum
C:Species: Clostridium acetobutylicum
C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 30-Sep-2001
C:Accession: E96954
R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
J. Bacteriol. 183, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium *Clostridium acetobutylicum* ATCC824
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: E96954
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-393 <KUR>
A:Cross-references: GB:AE001437; PIDN:AAK78424.1; PID:gl5023300; GSPDB:GN00168
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:

A:Gene: CAC0444
C:Superfamily: Aquifex aeolicus Na+/H+-exchanging protein napA1

Query Match 5.5%; Score 103.5; DB 2; Length 393;
Best Local Similarity 22.4%; Pred. No. 0.28;
Matches 82; Conservative 59; Mismatches 114; Indels 111; Gaps 21;

Qy 19 MVQNHAD--MVSCVGMFFVLGLMFECTAEMSI-----VFLTQLQGVVVPVPAEGLPSG 67
Db 47 LIQANADIKLLSLGVVFF---LMLAGIETNLDLDELKAGKSSFLIALGGLIIP---LIVG 100
Qy 68 SRLYHYGVKDLATVFFYMLVAIIHATIOEYVLDKLSRRLQLTGKONKLEAGQLSVF 127
Db 101 TLSAYMF---FSNFYENLFFVGVILTAT-----SVSISVQ--TLTEGLKLRSGINIL 148
Qy 128 --YIVSGINGMIILASENCLSPTLLMKSQPHNMTFMKFFEYI-----SOLAYWFHSHF 179
Db 149 GAATIDDLGLILITVLAISGCT---KSHGSSIF-MTFIYIGIFCLVSLAIAFLPK 202
Qy 180 P-ELYFQKVRKQDIPQOLIYIGLHLPHIG-----GAYLLYLNHLLGULL 220
Db 203 PIDKLTQKFKPKR-----GLAIFSAALICAFATKLGIAAIAITGAYI-----CGLV 249
Qy 221 LLMHLY-----AVELLS-SVCSLLYF---GDERYQKGLSLWPIVFISGRLYVTLIVSVVT 270
Db 250 LSPITHKEYIEKRVKIISTSFSPFASVGSASVKGILNFEVL-----LITLMFIIA 303
Qy 271 VGLHLAGTNRN-----GNALSGNVNVLAAKIA-----VLSSCSQIYVIT 310
Db 304 VICKILGCSASALTFLFKKSEALQIGVMVSRGEVAIITANIGLOAKIIEEITFLPTLIV 363
Qy 311 WTLTVV 316
Db 364 VILTVV 369

RESULT 12
B83987
ABC transporter (permease) B82698 [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: B83987
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Fujii, F.; I
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* a
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: B83987
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-522 <STO>
A:Cross-references: GB:AP001516; GB:BA000004; NID:gi0175192; PIDN:BA06417.1; GSPDB:G
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH2698

Query Match 5.4%; Score 102.5; DB 2; Length 522;
Best Local Similarity 23.2%; Pred. No. 0.48;
Matches 67; Conservative 57; Mismatches 66; Indels 99; Gaps 21;

Qy 14 LSHE-FMVQNHADMVSCVGMFFVLGLMFECTAEM-----SIVFLTQLQGVVVPVPAEGLPS 66
Db 75 LSHGIFMKQRKEF-----GLFSLGMSYQDMYKMLLENAGIAFLSL-----VVGULLS 123
Qy 67 GS-----RTLYHYGVKDLATVFF---FYMLVAIIHATIOEYVLDKLSRRLQLTGK-- 113
Db 124 GTVFSLRFLFSTMYIEVDIDISFQLOCTFY-LYATAFSVI--FIL-AIGKTLFTIRGOK 179
Qy 114'-----KONKLEAGQLSVFIVSGINGMIILASENCLSPTLLMKSQPHNMTFMKFFEY 168
Db 180 IIVAKMENRLSEQKQSPW--LGGIGGILVWAS-----LVFLY 215

QY 169 ISQLAYWHSPEL--YFQVKRKQDIPQ--LIYIGLHFGHGGAYLLYLNHGLLLML 224
DB 216 VNAI-----DOYPMDDGYF-----LIGSTVFLIGY-----LALSQSGSLIQM 255
QY 225 HYAVELLSSVC--SLLYGDERYOKGLSLWPVIFISGRVTLIVSVTV 271
DB 256 ---VKNPSIYRNLLHLSNLNY-KFKQLTSIFL-----LIVMTWTI 295

RESULT 13
H86268
hypothetical protein F1384.7 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2001
C:Accession: H86268
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;
ansen, C.W.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maitl, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719
A:Accession: H86268
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-308 <STO>
A:Cross-references: GB:AE005172; NID:99802756; PIDN:AAF99825.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1

Query Match 5.3%; Score 101; DB 2; Length 308;
Best Local Similarity 21.8%; Pred. No. 0.35;
Matches 42; Conservative 37; Mismatches 84; Indels 30; Gaps 6;
QY 74 GVKDLATVFFYMLVAIIHATIQEYVLDKLSRLQLTGKQN-----KLANE 120
DB 23 FVLPLEAVFPPS-----IRELDRFVEKLAIVYKHKRQMDGDTTERKKKIRKE 77
QY 121 AGQLSVFIVSGWGMILASENCLSDPTLW-----KSOPHNMFMOMKFFYISQLAY 175
DB 78 SAWKCVYVLSAEILALSYNEPMFMNTKYFWGPGDQTPDQOTKLKLLYMFVAGFY 137
QY 176 FHSFPPELYFQVKRKQDIPGQLIYIGLH--LFIHGGAYLLYLNHGLLLMLHYAVELL 232
DB 138 TYSIFALFWETRSDFG---VSMGHHTATLILIVSVCSFSRVSGLVLDHSDVFL 194
QY 233 SVCSL-LYFGDER 244
DB 195 EVGKMSKYSGAER 207

RESULT 14
T27324
hypothetical protein Y683B.10 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T27324
R:White, S.
submitted to the EMBL Data Library, October 1998
A:Reference number: Z20345
A:Accession: T27324
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-360 <WIL>
A:Cross-references: EMBL:AL032655; PIDN:CAA21723.1; GSPDB:GN00019; CESP:Y683B.10
A:Experimental source: clone Y683B
C:Genetics:
A:Gene: CESP:Y683B.10

A:Map position: 1
A:Introns: 24/3; 58/1; 88/3; 143/1; 205/2; 310/3

Query Match 5.3%; Score 100.5; DB 2; Length 360;
Best Local Similarity 20.0%; Pred. No. 0.46;
Matches 41; Conservative 46; Mismatches 89; Indels 29; Gaps 7;

QY 87 LVATIIHATIQEYVLDKLSRLQLTGKQN-----KLANEAGQLSVFVIVSG 132
DB 66 LYTVLILASITFRLYVLIQIRLESWTQOHNIYPRFAHKVPSEFWKLTYYGTWIF----A 121
QY 133 IWGMILILASENCLSDPTLL--WKSOHNMFMOMKFFYISQLAYWPHSPPELYFQVKRK 189
DB 122 FYFMCVDSHDIFNDPLSMWIEWGSGRPMHQVQVIYAVQSAFYIHSIYATLFDLWR 181
QY 190 QDIPQQLIYIGLHFGHGGAYLLYLNH---GULLMLHYAVELLSSVCSLLYFGDERYO 246
DB 182 KD--SWLMFVH-HFIALGLLFLSYVDNFTLPGALVFLHDNSDATLEITKLSFLYKRTN 238
QY 247 KGLSIPWIFISGRVTLIVSVTV 271
DB 239 R--QYKYFPLMGNAFILFALLVW 261

RESULT 15
D71467
probable tyrosine transport - Chlamydia trachomatis (serotype D, strain UW3/Cx)
C:Species: Chlamydia trachomatis
C:Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 08-Oct-1999
C:Accession: D71467
R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitche
Science 282, 754-759, 1998
A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia t
A:Reference number: A71570; MUID:99000809
A:Accession: D71467
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-397 <ARN>
A:Cross-references: GB:AE001354; GB:AE001273; NID:93329280; PIDN:AAC68415.1; PID:9332
A:Experimental source: serotype D, strain UW-3/Cx
C:Genetics:
A:Gene: typ_P_2
C:Superfamily: tyrosine-specific transport protein

Query Match 5.3%; Score 100.5; DB 2; Length 397;
Best Local Similarity 21.1%; Pred. No. 0.52;
Matches 64; Conservative 43; Mismatches 109; Indels 87; Gaps 14;

QY 12 PVLSEFVONHADVMSCVGMFFVGLM-----FEGTAEMSIYVFL 53
DB 138 PLLMANTSVIDYCNRGFVGLIFVGLLVLPRIQGEILLRASFSSLSLPFFLAF 197
QY 54 QHGVVPAEGLPSGRTLYH--GVKDL-ATVFFYMLVAIIHATIQEYVLDKLSRLQ 109
DB 198 GFQNVVP-----SLYHLDGNIREVKRAILGLSLPILYIAEALVLVTP-LVD 247
QY 110 LTKGQKNKLANEAGOL-----SVFYIVSGIWMILLASE-----NCLSD---PTLLWKSQ 155
DB 248 LLKAKDGLGTAAGALQGLSLNSAFYINAGELFGFALVTSFGTALAKDPIYIDIFKWDAR 307
QY 156 PHNMTFMOMKFFYISQL--AYWPHSPPELYFQVKRKQDIPGQLIYIGLHFGHGGAYLLY 213
DB 308 KK-----RVSLFFLVQVFPVLPVWAIIFYPEIVLSCLR-----YAG---GIGGACIIV 349
QY 214 LNHGLGLLLMLHYAVELLSSVCSLLYFGDERYOKGLSLWPVIFISGRVTLIVSVTV 273
DB 350 LFPVAML-----WNGRYKRRCFGRKILPGGKTVLLILTGV-L 368
QY 274 HLA 276
DB 389 NLA 391

Search completed: September 20, 2002, 14:36:27
Job time: 139 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2002, 14:35:13 ; Search time 13.4 Seconds
(without alignments)
1048.895 Million cell updates/sec

Title: us-09-807-470-2

Perfect score: 1892

Sequence: 1 MGLRKNRNPVLSHEFMV.....NGVENPNRIDSPPKKKRP 363

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1145.5	60.5	373	1 TRAM_CANFA	Q01685 canis famil
2	1129.5	59.7	373	1 TRAM_HUMAN	Q15629 homo sapien
3	1084.5	57.3	358	1 TRAM_BOVIN	Q09k24 bos taurus
4	754	39.9	370	1 Y557_HUMAN	Q15035 bos sapien
5	148.5	7.8	411	1 LAG1_YEAST	P38703 saccharomyc
6	137	7.2	418	1 YKAB_YEAST	P28496 saccharomyc
7	133	7.0	384	1 YHFX_SCHPO	O59735 schizosacch
8	125	6.6	390	1 LAG1_SCHPO	P78970 schizosacch
9	97.5	5.2	533	1 GLPT_HUMAN	P57057 homo sapien
10	97	5.1	1450	1 CLTR_HUMAN	Q00554 oryctolagus
11	96.5	5.1	534	1 COX1_RICPR	O54069 rickettsia
12	95	5.0	370	1 YDIK_ECOLI	P77175 escherichia
13	94.5	5.0	540	1 SOAL_MOUSE	O61263 mus musculus
14	94.5	5.0	788	1 QOXM_SULAC	P39481 sulfolobus
15	94	5.0	272	1 ATP6_BUCAP	O51878 buchnera ap
16	93.5	4.9	521	1 COX1_APILI	P20374 apis mellif
17	93	4.9	532	1 YHCA_BACSU	P54585 bacillus su
18	91.5	4.8	369	1 Y316_MYCPN	O50361 mycoplasma
19	91	4.8	307	1 OGD2_HUMAN	P58182 homo sapien
20	91	4.8	624	1 NKX3_RAT	O9epq0 ratus norv
21	91	4.8	645	1 NKX3_MOUSE	O99p07 mus musculus
22	90	4.8	513	1 COX1_RABIT	O79429 oryctolagus
23	90	4.8	514	1 COX1_CERSI	O03198 ceratotheri
24	90	4.8	514	1 COX1_EQUAS	P92477 equus asinu
25	90	4.8	514	1 COX1_HORSE	P48659 equus cabal
26	90	4.8	514	1 COX1_RHJUN	O96052 rhinoceros
27	90	4.8	516	1 COX1_GADMO	Q36775 gadus morhu
28	89	4.7	360	1 MRAY_PASMO	P57816 pasteurella
29	89	4.7	430	1 SECY_BACHD	P38375 bacillus ha
30	89	4.7	499	1 MVIN_AQUAE	O67658 aquifex aeo
31	89	4.7	514	1 COX1_BOVIN	P00396 bos taurus
32	89	4.7	514	1 COX1_CANFA	Q92264 canis famil
33	89	4.7	514	1 COX1_FELCA	P48888 felis silve

ALIGNMENTS

RESULT 1

ID	TRAM_CANFA	STANDARD;	PRT;	373 AA.
AC	Q01685;			
DT	01-JUL-1993 (Rel. 26, Created)			
DT	01-JUL-1993 (Rel. 26, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	TRAM protein (Translocating chain-associating membrane protein).			
OS	Canis familiaris (Dog).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.			
OX	NCBI_TaxID=9615;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 1-26 AND 164-184.			
RC	TISSUE=Kidney;			
RX	MEDLINE=92244357; PubMed=1315422;			
RA	Goerlich D., Hartmann E., Prehn S., Rapoport T.A.;			
RT	"A protein of the endoplasmic reticulum involved early in polypeptide translocation."			
RL	Nature 357:47-52(1992).			
CC	!- FUNCTION: STIMULATORY OR REQUIRED FOR THE TRANSLOCATION OF			
CC	SECRETORY PROTEINS ACROSS THE ER MEMBRANE.			
CC	!- SUBCELLULAR LOCATION: Endoplasmic reticulum membrane.			
CC	!- SIMILARITY: BELONGS TO THE LASSI FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; X63678; CAA45217.1; ..			
DR	PIR; S21736; S21736.			
KW	Endoplasmic reticulum; Transmembrane; Glycoprotein; Translocation.			
FT	INIT_MET 0			
FT	DOMAIN 1 28			
FT	TRANSMEM 29 49			
FT	DOMAIN 50 75			
FT	TRANSMEM 76 96			
FT	DOMAIN 97 120			
FT	TRANSMEM 121 141			
FT	DOMAIN 142 158			
FT	TRANSMEM 159 179			
FT	DOMAIN 180 191			
FT	TRANSMEM 192 212			
FT	DOMAIN 213 216			
FT	TRANSMEM 217 237			
FT	DOMAIN 238 250			
FT	TRANSMEM 251 271			
FT	DOMAIN 272 296			
FT	TRANSMEM 297 317			
FT	DOMAIN 318 373			
FT	CARBOHYD 35 55			
SQ	SEQUENCE 373 AA; 43029 MW; ID85808E1D80E835 CRC64;			

078749 ovis aries
Q9hc58 homo sapien
O00337 homo sapien
P35071 bos taurus
Q00555 ovis aries
Q9zc55 rickettsia
Q92zy9 hippopotamu
Q10487 schizosacch
Q04443 bacillus fi
P18943 gallus gall
O99191 salmonella
O79876 sus scrofa

Query Match 60.5%; Score 1145.5; DB 1; Length 373;
 Best Local Similarity 59.8%; Pred. No. 2.7e-88;
 Matches 222; Conservative 58; Mismatches 80; Indels 11; Gaps 4;

QY 3 LRKKARNPPVLSHEFWQNHADVSCVGMFFVLGLMFEQTAEMSIIVFLTQHGQVVPVPA- 61
 DB 2 IRKSKTSPVLSHEFVILQNHADIVSCVAMVFLGLMFEITAKASIIFFVTLQVNVLPAT 61

QY 62 EQLSGSRTLYHYGVKDLATVFFVYMLVAIIHATIQEYVLDKLSRLQLTGKQNKLEA 121
 DB 62 EQATESVSLYYIGIKDLATVFFVYMLVAIIHATIQEYVLDKLSRLQLTGKQNKLEA 121

QY 122 GOLSVFYIVSGWGMILASNCUSDPTLLWKQPHNMFTQMKFFYISOLAYWFFHSFPE 181
 DB 122 GOLSAFYILFCWGTFFILISENYISDPTILWRAYPHNLMFTQMKFFYISOLAYWFFHAFFE 181

QY 182 LYFQKVRKQDIPGOLIVIGLHPLHIGAGYLLYNHLGLLLMLHYAVELLSVCSLLYFG 241
 DB 182 LYFQKVRKQDIPGOLIVIGLHPLHIGAGYLLYNHLGLLLMLHYAVELLSVCSLLYFG 241

QY 242 DERYOKGLSLWPIVIFISGRVTLIVSVTVGLHLA-GTNRNGNALSGNVNVAIAKIAVL 300
 DB 242 DERYOKGLSLWPIVIFISGRVTLIVSVTVGLHLA-GTNRNGNALSGNVNVAIAKIAVL 300

QY 301 SSCSIQYIITWTLTWTWVWLRWLEDANLHVCGRKRR-----SRSRKGTEGVVE---NPNR 351
 DB 301 SSCSIQYIITWTLTWTWVWLRWLEDANLHVCGRKRR-----SRSRKGTEGVVE---NPNR 351

QY 352 IDSPPKKKEKA 362
 DB 352 IDSPPKKKEKA 362

QY 362 ADSPRNKKEKS 372
 DB 362 ADSPRNKKEKS 372

RESULT 2
 TRAM_HUMAN
 ID TRAM_HUMAN STANDARD; PRT; 373 AA.
 AC Q15629;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE TRAM protein (translocating chain-associating membrane protein).
 GN TRAM.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92244357; PubMed=1315422;
 RA Goerlich D., Hartmann E., Prehn S., Rapoport T.A.;
 RA Strausberg R.;
 RC TISSUE=Kidney;
 RL Nature 357:47-52(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 RA Strausberg R.;
 RC TISSUE=Kidney;
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: STIMULATORY OR REQUIRED FOR THE TRANSLOCATION OF
 CC SECRETORY PROTEINS ACROSS THE ER MEMBRANE.
 CC -1- SUBCELLULAR LOCATION: Endoplasmic reticulum membrane.
 CC -1- SIMILARITY: BELONGS TO THE LASS1 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; X63679; CAA45218.1; --
 DR EMBL; BC000687; AAH00687.1; --
 KW MIM; 605190; --
 FT Endoplasmic reticulum; Transmembrane; Glycoprotein; Translocation.
 FT INIT_MET 0 0
 FT DOMAIN 1 28 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 29 49 POTENTIAL.
 FT DOMAIN 50 75 LUMENAL (POTENTIAL).
 FT TRANSMEM 76 96 POTENTIAL.
 FT DOMAIN 97 120 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 121 141 POTENTIAL.
 FT DOMAIN 142 158 LUMENAL (POTENTIAL).
 FT TRANSMEM 159 179 POTENTIAL.
 FT DOMAIN 180 191 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 192 212 POTENTIAL.
 FT DOMAIN 213 216 LUMENAL (POTENTIAL).
 FT TRANSMEM 217 237 POTENTIAL.
 FT DOMAIN 238 250 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 251 271 POTENTIAL.
 FT DOMAIN 272 296 LUMENAL (POTENTIAL).
 FT TRANSMEM 297 317 POTENTIAL.
 FT DOMAIN 318 373 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 55 55 N-LINKED (GLCNAC...) (PROBABLE).
 SQ SEQUENCE 373 AA; 42940 MW; C220949AF4EFDD0 CRC64;

Query Match 59.7%; Score 1129.5; DB 1; Length 373;
 Best Local Similarity 58.8%; Pred. No. 5.8e-87;
 Matches 218; Conservative 61; Mismatches 81; Indels 11; Gaps 4;

QY 3 LRKKARNPPVLSHEFWQNHADVSCVGMFFVLGLMFEQTAEMSIIVFLTQHGQVVPVPA- 61
 DB 2 IRKSKTSPVLSHEFVILQNHADIVSCVAMVFLGLMFEITAKASIIFFVTLQVNVLPAT 61

QY 62 EQLSGSRTLYHYGVKDLATVFFVYMLVAIIHATIQEYVLDKLSRLQLTGKQNKLEA 121
 DB 62 EQATESVSLYYIGIKDLATVFFVYMLVAIIHATIQEYVLDKLSRLQLTGKQNKLEA 121

QY 122 GOLSVFYIVSGWGMILASNCUSDPTLLWKQPHNMFTQMKFFYISOLAYWFFHSFPE 181
 DB 122 GOLSAFYILFCWGTFFILISENYISDPTILWRAYPHNLMFTQMKFFYISOLAYWFFHAFFE 181

QY 182 LYFQKVRKQDIPGOLIVIGLHPLHIGAGYLLYNHLGLLLMLHYAVELLSVCSLLYFG 241
 DB 182 LYFQKVRKQDIPGOLIVIGLHPLHIGAGYLLYNHLGLLLMLHYAVELLSVCSLLYFG 241

QY 242 DERYOKGLSLWPIVIFISGRVTLIVSVTVGLHLA-GTNRNGNALSGNVNVAIAKIAVL 300
 DB 242 DERYOKGLSLWPIVIFISGRVTLIVSVTVGLHLA-GTNRNGNALSGNVNVAIAKIAVL 300

QY 301 SSCSIQYIITWTLTWTWVWLRWLEDANLHVCGRKRR-----SRSRKGTEGVVE---NPNR 351
 DB 301 SSCSIQYIITWTLTWTWVWLRWLEDANLHVCGRKRR-----SRSRKGTEGVVE---NPNR 351

QY 352 IDSPPKKKEKA 362
 DB 352 IDSPPKKKEKA 362

QY 362 ADSPRNKKEKS 372
 DB 362 ADSPRNKKEKS 372

RESULT 3
 TRAM_BOVIN
 ID TRAM_BOVIN STANDARD; PRT; 358 AA.
 AC Q9GK24;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE TRAM protein (translocating chain-associating membrane protein).
 GN TRAM.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.

OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Clark T.G., Morris J., Akamatsu M., McGraw R.A., Ivarie R.D.;
 RT "Cloning and sequence analysis of a bovine tran cDNA";
 RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: STIMULATORY OR REQUIRED FOR THE TRANSLATION OF
 CC SECRETORY PROTEINS ACROSS THE ER MEMBRANE (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM MEMBRANE (By
 CC similarity).
 CC -!- SIMILARITY: BELONGS TO THE LASS1 FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: U19578; AAG10391.1; -
 KW Endoplasmic reticulum; Transmembrane; Glycoprotein; Translocation.
 FT NON_TER 1
 FT DOMAIN <1 13 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 14 34 POTENTIAL.
 FT DOMAIN 35 60 LUMENAL (POTENTIAL).
 FT TRANSMEM 61 81 POTENTIAL.
 FT DOMAIN 82 105 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 106 126 POTENTIAL.
 FT DOMAIN 127 143 LUMENAL (POTENTIAL).
 FT TRANSMEM 144 164 POTENTIAL.
 FT DOMAIN 165 176 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 177 197 POTENTIAL.
 FT DOMAIN 198 201 LUMENAL (POTENTIAL).
 FT TRANSMEM 202 222 POTENTIAL.
 FT DOMAIN 223 235 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 236 256 POTENTIAL.
 FT DOMAIN 257 281 LUMENAL (POTENTIAL).
 FT TRANSMEM 282 302 POTENTIAL.
 FT DOMAIN 303 358 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 40 40 N-LINKED (GLCNAC...) (PROBABLE).
 SQ SEQUENCE 358 AA; 41403 MW; 25CF9930C4CDDA15 CRC64;
 Query Match 57.3%; Score 1084.5; DB 1; Length 358;
 Best Local Similarity 59.9%; Pred. No. 3.1e-83;
 Matches 214; Conservative 53; Mismatches 79; Indels 11; Gaps 5;
 QY 17 EFWQNHADNVSCVGMFFVLGLMFEPTAEMSIIVFTLQHGVVVPA-BGLPSGSRITLHYG 75
 DB 1 EFWLQNHADIVSCVAMVFLGLMFEPTAKVSIIVFTLQYNTLPATEQATSAFLYYG 60
 QY 76 VKDLATVFFMLVAIIHAIIOEYVLDKLSRRLQITKQNKLEAGOLSVFIIVSGIWG 135
 DB 61 IKDLATVFFMLVAIIHAIIOEYVLDKINRMHFSKTKHFNESGOLSAFLFSCIWG 120
 QY 136 MTLASENCLSDPTLLKWSQPHNMFTQMKFFYSQLAYWPHSPPELYFQKVRKQDIPGQ 195
 DB 121 TFLISENVISDPTILRAYPHNLTQMKFFYSQLAYWPHAFPELYFQTKKEDIPRO 180
 QY 196 LIYIGLHFIHGAYLLNLHGLLLMLHYAVELLSCVLSLFGDERYOKGLSLMPIV 255
 DB 181 LYVIGLYLFIHAGAYLLNLHGLVLLVHYVEFLEHISRLFYFSDKEYOKGFSWLAVL 240
 QY 256 FTSGRLVTLIVSVTVVGLHLA-GTNRNGNALSGNNVLAALKIAVLSSCSCTOVYITWTLT 314
 DB 241 FVLGRLETLILSVLGVGLARAENOKLDFSTGNFNLAVRIAVLASICITQAFMMKKFI 300
 QY 315 TVWLQWLEDANLHVCGKRK-----RSRS-RKGTENGVE---NPNRIDSPPKKKEA 362
 DB 301 NFOLRRWREHSAFOAPAVKKPPVTKGRSXXKGTENGVTSGNAGDSRNRKES 357

RESULT 4
 Y557_HUMAN
 ID Y557_HUMAN STANDARD; PRT; 370 AA.
 AC Q15035;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein KIAA0057.
 GN KIAA0057.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Bone marrow;
 RX MEDLINE=96051398; PubMed=7584044;
 RA Nomura N., Nagase T., Miyajima N., Sazuka T., Tanaka A., Sato S.,
 RA Seki N., Kawarabayashi Y., Ishikawa K.-I., Tabata S.;
 RT "Prediction of the coding sequences of unidentified human genes. II.
 RT The coding sequences of 40 new genes (KIAA0041-KIAA0080) deduced by
 RT analysis of cDNA clones from human cell line KG-1.";
 RL DNA Res. 1:223-229(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Tracey A.;
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -!- SIMILARITY: BELONGS TO THE LASS1 FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: D31762; BAA06540.1; -
 DR EMBL: AL049611; CAB71119.1; -
 KW Hypothetical protein; Transmembrane.
 FT TRANSMEM 23 43 POTENTIAL.
 FT TRANSMEM 76 96 POTENTIAL.
 FT TRANSMEM 120 140 POTENTIAL.
 FT TRANSMEM 160 180 POTENTIAL.
 FT TRANSMEM 199 219 POTENTIAL.
 FT TRANSMEM 251 271 POTENTIAL.
 FT TRANSMEM 288 308 POTENTIAL.
 SQ SEQUENCE 370 AA; 43327 MW; 9B5183F1A3D45366 CRC64;
 Query Match 39.9%; Score 754; DB 1; Length 370;
 Best Local Similarity 42.4%; Pred. No. 1.1e-55;
 Matches 158; Conservative 63; Mismatches 132; Indels 20; Gaps 5;
 QY 5 KKNARNPVLVSHEFVNQADHVMSCVGMFFVLGLMFEPTAEMSIIVFTLQHGVVVPAEGL 64
 DB 4 RRTKSYPLFSQEFVIHNDHAGIDGFCVLCVGLMFEVTAFTAFILPOYNISVPT 60
 QY 65 PSGSRITL-YHGVKDLATVFFMLVAIIHAIIOEYVLDKLSRRLQITKQNKLEAGOL 123
 DB 61 -ADSETVHYHGPDLVTLFYIFITILLHVVQYIILDKISKRHLHSKVHSPNESGQ 119
 QY 124 LSVFIVSGINGMIIASLNCISDPTLLKWSQPHNMFTQMKFFYSQLAYWPHSPPELY 183
 DB 120 LVVFHTSVINCFFVVVTEGYLTNPSRWEDYPHVLPFQVQKFFYLCOLAYWLHALPELY 179
 QY 184 FOKVRKQDIPGOLYIGLHFIHGAYLLNLHGLLLMLHYAVELLSCVLSLFGDERY 243
 DB 180 FOKVRKEETPROLYICLYLVHAGAYLLNLHGLVLLVHYVEFLEHISRLFYFSDKEY 239
 QY 244 RYQKGLSLWPIVIFISGRVTLIVSVTVVGLHLA-GTNRNGNALSGNNVLAALKIAVLSSS 302

DB 240 NNEKLSAWAAYGVGTRFLITLAVLAIGFLARMENQAFDEPKGNFTLFCRLCVLLIV 299
 QY 303 CSIQVYITTLTTLVQLRWLEDANLHVCGRKR-----SRSKGTGVENPEN 350
 DB 300 CAAQWLWREHISQLRHWRYNEQ--SAKRRVPATPRLPARLIKRESYHGVVKA 357
 QY 351 RIDSPKKKAP 363
 DB 358 NGTSPTKKLSP 370

RESULT 5
 LAG1_YEAST
 ID LAG1_YEAST STANDARD; PRT; 411 AA.
 AC P38703;
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Longevity-assurance protein 1 (Longevity assurance factor 1).
 GN LAG1 OR YHL003C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-X2180-1A;
 RX MEDLINE=94253121; PubMed=8195187;
 RA D'Mello N.P., Childress A.M., Franklin D.S., Kale S.P.,
 RA Plinswasdi C., Jazwinski S.M.;
 RA "Cloning and characterization of LAG1, a longevity-assurance gene in
 RT yeast.";
 RL J. Biol. Chem. 269:15451-15459(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-S288C / AB972;
 RX MEDLINE=94378003; PubMed=8091229;
 RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,
 RA Du Z., Favell A., Fulton L., Gattung S., Geisel C., Kirsten J.,
 RA Kucaba T., Hillier L., Jier M., Johnston L., Langston Y.,
 RA Latreille P., Louis E.J., Macris C., Mardis E., Menezes S., Mouser L.,
 RA Nham M., Rifkin L., Riles L., St Peter H., Trevaskis E., Vaughan K.,
 RA Vignati D., Wilcox L., Wohlman P., Waterston R., Wilson R.,
 RA Vaudin M.;
 RA "Complete nucleotide sequence of Saccharomycetes cerevisiae chromosome
 RT VIII.";
 RL Science 265:2077-2082(1994).
 CC -1- FUNCTION: INVOLVED IN THE AGING PROCESS. DELETION OF LAG1 RESULTS
 CC IN A PRONOUNCED INCREASE (APPROXIMATELY 50%) IN MEAN AND IN
 CC MAXIMUM LIFE SPAN.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: BELONGS TO THE LASSI FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U08133; AAA21579.1; -
 DR EMBL; U10555; AAB68429.1; -
 DR PIR; S46800; S46800.
 DR SGD; S0000995; LAG1.
 DR Transmembrane.
 KW TRANSNEM 82 102 POTENTIAL.
 FT TRANSNEM 135 155 POTENTIAL.
 FT TRANSNEM 177 197 POTENTIAL.
 FT TRANSNEM 212 232 POTENTIAL.
 FT TRANSNEM 252 272 POTENTIAL.
 FT TRANSNEM 297 317 POTENTIAL.
 FT TRANSNEM 356 376 POTENTIAL.

FT CONFLICT 173 174 ML -> IV (IN REF. 2).
 FT CONFLICT 220 220 F -> C (IN REF. 2).
 FT CONFLICT 301 411 VFTPFVGLFVFFWVIRLVNIRILSVILSVLTFRHEGNVVL
 FT NEATQYKWCWISPIVFEVLAALQVNLVWLFLILRLVRL
 FT IMQGIQKDSRSDSDSAENESKEKE -> TEISGIME
 FT KQSIDNDNPTERALSNETSKOVKPDLLVLPNTENRAL
 FT LEAKRSVPTIADTDSEPSLVYPIPGNDOSLSRVNELL
 FT GVLARAGQRLQRLARNNEK (IN REF. 1).
 SQ SEQUENCE 411 AA; 48454 MW; 91676D56AC053F3C CRC64;

Query Match 7.8%; Score 148.5; DB 1; Length 411;
 Best Local Similarity 22.0%; Pred. No. 3.7e-05;
 Matches 68; Conservative 46; Mismatches 106; Indels 89; Gaps 11;

QY 26 MVSQVGMFFVLGLMFEGETAEMSVFLTLQHGVSVPVPAEGLSPGSGRTLYHYGVKDLATVFFY 85
 DB 91 LVCVYSAYFLSGNRTESNPLHMFVAISYQ-----VDGTD-----YAKGIKDLSEVFFY 139
 QY 86 MLVAIIITHATIQEVVLDKLSRRRLQTLTKGKQNKLEAGQLSVFIY-VSGIWMIL-ASEN 143
 DB 140 MIPTFLREFLMDVIRPFTVYLVNVTSEHQRKRMLEQMYAIFYGVSGPGLYIMVHSDL 199
 QY 144 CLSDPTLLMKSQPHNMFTQMKFFYISQLAYW-----FHSFPE 181
 DB 200 WLEKTKPMYRTPVITNPFLEKIFYLQAAFAWAQACVVLQLEKPKDYKELVFHHIVT 259
 QY 182 L-----YFQVKRKQIDIPGQLIYI-----GLHLFH 205
 DB 260 LLLIWSYVFHTKM-----GLAIYITMDVSDFFLSLTKLYNSVTFPPVFGFLVF- 312
 QY 206 IGGAYLLYLNHLGLLLHLHYAVELLSSVCSLLYFGDERYQKGLSLWPIVFIISGLRVTLI 265
 DB 313 ----FWILRHVVNIRILSVLTFRHEGNVNLNFATQYKCKWISL-PIVFEV-----LI 361
 QY 266 VSVTVTVGLH 274
 DB 362 AALQLVNLV 370

RESULT 6
 YKAB_YEAST
 ID YKAB_YEAST STANDARD; PRT; 418 AA.
 AC P28496;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical 49.0 kDa protein in UFD4-CAP1 intergenic region.
 GN YKL008C OR YKL156
 OS Saccharomycetes cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93255906; PubMed=8498728;
 RA Boyer J., Pascolo S., Richard G.F., Dujon B.;
 RT "Sequence of a 7.8 kb segment on the left arm of yeast chromosome XI
 RT reveals four open reading frames, including the CAP1 gene, an intron-
 RT containing gene and a gene encoding a homolog to the mammalian UOG-1
 RT gene.";
 RL Yeast 9:279-287(1993).
 RN [2]
 RP SEQUENCE OF 1-149 FROM N.A.
 RX MEDLINE=93077675; PubMed=1447293;
 RA Amatruda J.F., Gattermeir D.J., Karpova T.S., Cooper J.A.;
 RT "Effects of null mutations and overexpression of capping protein on
 RT morphogenesis, actin distribution and polarized secretion in yeast.";
 RL J. Cell Biol. 119:1151-1162(1992).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: BELONGS TO THE LASSI FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; X61398; CAA43670.1; -
 DR EMBL; S59773; AAC60549.1; -
 DR EMBL; Z28008; CAA81843.1; -
 DR PIR; S30134; S30134.
 DR SGD; S0001491; YKL008C.
 KW Hypothetical protein; Transmembrane.
 FT TRANSMEM 82 102 POTENTIAL.
 FT TRANSMEM 135 155 POTENTIAL.
 FT TRANSMEM 173 193 POTENTIAL.
 FT TRANSMEM 260 280 POTENTIAL.
 FT TRANSMEM 297 317 POTENTIAL.
 FT TRANSMEM 356 376 POTENTIAL.
 SQ SEQUENCE 418 AA; 48992 MW; 7691BA623AC0460A CRC64;

Query Match 7.2%; Score 137; DB 1; Length 418;
 Best Local Similarity 23.5%; Pred. No. 0.00034;
 Matches 61; Conservative 44; Mismatches 93; Indels 62; Gaps 11;

QY 67 GSTRTHYGVKDLATVFFVYMLVAIIHATIQEYVLDKLSRRQLTKGQKNLNEAGOLSV 126
 DB 121 GDTNAYGKGINLDFVFFYVFFTEFLREFLMDVIRPFAIRLHVTSKRIKRIEOMYAI 180
 QY 127 FYI-VSGIWMIL-ASENCLESDPTLLMKSPHNMFTQMKFFYSOLAYW----- 175
 DB 181 FYTVSGPGIYCMYHSDLWFFNTAMRYTDFTNPLFKFYLGQAQFAWAQACILVL 240
 QY 176 -----FHSFPEL-----YFKVRKQDIPGQLIYIGLHLPHIGGAYLLXNH 216
 DB 241 QLEKPKRKHNETFHHTVLLWSSVYFHTKMGLP---IVITMDVSDFLSFKTLNY 297
 QY 217 L-GLL-----LMLHYA-VELLSSVCS-----LIYFGDERYQKGLSLWPI 254
 DB 298 LSGLAFFSFAIVVAMIVYLRHYINLKILMSVLTOFTREGNYVLNFAQQYKWCWISL-PI 356
 QY 255 VFISGRVTLIVSVTVGLH 274
 DB 357 VFV-----LIGALQVLNLY 370

RESULT 7
 ID YHFX SCHPO STANDARD; PRT; 384 AA.
 AC O59735;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE Hypothetical 42.4 kDa protein C3E17.15c in chromosome II.
 GN SPB3E7.15C OR SPB3AF6.02C.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RA Lyne M., Wood V., Rajandream M.A., Barrell B.G., Brown D.,
 RA Churcher C.M.;
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RA William R., Rajandream M.A., Barrell B.G., Skelton J., Churcher C.M.;
 RA Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).

CC -!- SIMILARITY: BELONGS TO THE LASSI FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; AL023534; CAA19018.2; -
 DR EMBL; AL031534; CAA20722.2; -
 KW Hypothetical protein; Transmembrane.
 FT TRANSMEM 59 79 POTENTIAL.
 FT TRANSMEM 104 124 POTENTIAL.
 FT TRANSMEM 155 175 POTENTIAL.
 FT TRANSMEM 195 215 POTENTIAL.
 FT TRANSMEM 233 253 POTENTIAL.
 FT TRANSMEM 270 290 POTENTIAL.
 FT TRANSMEM 329 349 POTENTIAL.
 SQ SEQUENCE 384 AA; 45335 MW; 82FCF8EA6638849A CRC64;

Query Match 7.0%; Score 133; DB 1; Length 384;
 Best Local Similarity 20.5%; Pred. No. 0.00067;
 Matches 48; Conservative 45; Mismatches 117; Indels 24; Gaps 7;

QY 41 EGTAEISVFLTL-----OHGVVPAEGLSPGSRSLYHYGVKDLATVFFYML 87
 DB 56 EKTWIVPLILLTLGVWFPNPNVNGYKIFL-SYIPGTNPAQYKGRDLAFCLFYAL 114
 QY 88 VALIIHATQEVLDKLSRRLQ-LTKGKONLNEAGOLSVFYVSGIWMIL-ASENCLE 145
 DB 115 FTFCTREFIMQETIIRHGFNRAPAKLRFEEQAYCTLYFTVMSGLYVKNQTPMWF 174
 QY 146 SDPTLLWKSQPHNMFTQMKFFYSOLAYWFSFPELYFQ--KVRKQDIPGQLIYIGLH 203
 DB 175 FNTDAFWEYEPHYHVGSPKAFYLIEAWYIQQALVLILQLEKPRK-DFKELVVHHITL 233
 QY 204 FHIGGAYLLYLNHLLMLHYAVELLSSVCSLLYFGDERYQKGLSLWPIVFI 257
 DB 234 LLIGLSYFFHTWIGLAVFITMDTSDIWLALSCL-----NYVNTVIVYPIEVI 282

RESULT 8
 ID LAG1 SCHPO STANDARD; PRT; 390 AA.
 AC P78970; O13860;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Longevity-assurance protein 1 (Longevity assurance factor 1).
 GN LAG1 OR SPAC1A6.09C.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Chanda E.R., Lingner C., Ko Z., Young P.G.;
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RA Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V., Wood V.;
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: INVOLVED IN THE AGING PROCESS (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -!- SIMILARITY: BELONGS TO THE LASSI FAMILY.
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -


```

Db      893 ADTLAAGLGRGLPLVFTLITVSKILHHKMLHSLV---QAPMSTLNTLKAGGILNRFSGD 949
QY      79 LA-----TVF-FYMLVAIIHA-----TIOEVLDK-----LSRRLQTRKGQ 115
Db      950 IAILDDLPLTIFDFIQILLIVGCAIAVWSVLOQYIFLATVPVIAAFTLLRAYFLHTSQ 1009
QY      116 NK-LNEAGQLSVP-YIVSGTGMILLASENCLSPTLLWKSQPNMHTFQMKFFYSOL- 172
Db      1010 LKQLESGRSPIFHLVTSLKGLWTLRA-----FGROPY-----FETLFKALNLH 1055
QY      173 -AVWFHSPFELYFOKVRKQDIPGGLIYGLHFIHGAYLLYL-----NHLGLLLLM-- 223
Db      1056 TANWF-----LYLSTLRFQMRIMFV---LFFIAVAFISILTTGSEGEVGIILLAM 1107
QY      224 -----LHYAV-----ELSSVCSLLYFGD-----ERY 245
Db      1108 NIMSTLQWANNSSIDVDLSMQSVRFPMFIDMPTAKRSIKSPSSNCQLSKVMIIENOH 1167
QY      246 QKGLSLWPIVIFISGRLYTL-----IVSVTVGLHLAGTNRNGNALSGN 288
Db      1168 VKDDWVP---SGQMTVVGKLTAKYIDSGNAILENISPSPQORVCLLGRTG---SGK 1220
QY      289 VNVLAAKIAYLSSCSCSIQV-YITWTLFTVYLQWR 321
Db      1221 STLLSAFRLLSGEIQIDGVSW--DSITLQOW 1252

RESULT 11
COX1L_RICPR STANDARD; PRT; 534 AA.
AC 054069;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable cytochrome c oxidase polypeptide I (EC 1.9.3.1) (Cytochrome
DE AA3 subunit 1).
GN CTAD OR COXA OR RP405.
OS Rickettsia prowazekii.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsia; Rickettsiae; Rickettsia.
OX NCBI_TaxID=782;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MADRID E;
RA Sacheritz T., Kurland C.G., Andersson S.G.E.;
RT "The bacterial origin of mitochondria inferred from a phylogenetic
RT analysis of the cytochrome b and cytochrome c oxidase I genes.";
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MADRID E;
RC MEDLINE=99039499; PubMed=9823893;
RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,
RA Sacheritz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RT "The genome sequence of Rickettsia prowazekii and the origin of
RT mitochondria.";
RL Nature 396:133-140(1998).
CC -1- FUNCTION: CYTOCHROME C OXIDASE IS THE COMPONENT OF THE RESPIRATORY
CC CHAIN THAT CATALYZES THE REDUCTION OF OXYGEN TO WATER. SUBUNITS 1-
CC 3 FORM THE FUNCTIONAL CORE OF THE ENZYME COMPLEX. CO I IS THE
CC CATALYTIC SUBUNIT OF THE ENZYME. ELECTRONS ORIGINATING IN
CC CYTOCHROME C ARE TRANSFERRED VIA THE COPPER A CENTER OF SUBUNIT 2
CC AND HEME A OF SUBUNIT 1 TO THE BIMETALLIC CENTER FORMED BY HEME A3
CC AND COPPER B (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome
CC c + 2 H(2)O.
CC -1- PATHWAY: TERMINAL STEP IN THE RESPIRATORY CHAIN.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO THE HEME-COPPER RESPIRATORY OXIDASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; Y13855; CA74167.1; -;
DR EMBL; AJ235271; CAAL4862.1; -;
DR HSSP; P98002; IARI.
DR InterPro; IPR000883; COX1.
DR Pfam; PF00115; COX1; 1.
DR PRINTS; PR01165; CYCOXIDASE1.
DR PROSITE; PS00077; COX1; 1.
KW Oxidoreductase; Heme; Copper; Transmembrane; Respiratory chain;
KW Complete proteome.
FT TRANSMEM 35 55 POTENTIAL.
FT TRANSMEM 76 96 POTENTIAL.
FT TRANSMEM 97 117 POTENTIAL.
FT TRANSMEM 120 140 POTENTIAL.
FT TRANSMEM 165 185 POTENTIAL.
FT TRANSMEM 202 222 POTENTIAL.
FT TRANSMEM 254 274 POTENTIAL.
FT TRANSMEM 286 306 POTENTIAL.
FT TRANSMEM 320 340 POTENTIAL.
FT TRANSMEM 357 377 POTENTIAL.
FT TRANSMEM 396 416 POTENTIAL.
FT TRANSMEM 433 453 POTENTIAL.
FT TRANSMEM 475 495 POTENTIAL.
FT METAL 81 81 IRON (HEME A) (PROBABLE).
FT METAL 260 260 COPPER B (PROBABLE).
FT METAL 264 264 COPPER B (PROBABLE).
FT METAL 309 309 COPPER B (PROBABLE).
FT METAL 310 310 COPPER B (PROBABLE).
FT METAL 395 395 IRON (HEME A3) (PROBABLE).
FT METAL 397 397 IRON (HEME A) (PROBABLE).
SQ SEQUENCE 534 AA; 59261 MW; 326685753548C3E CRC64;

Query Match 5.1%; Score 96.5; DB 1; Length 534;
Best Local Similarity 18.6%; Pred. No. 1.1;
Matches 65; Conservative 57; Mismatches 107; Indels 121; Gaps 19;
QY 14 LSHEPWONHADMVSCVM---FFVLGLMFEGTAEMSVFL-----TLQHGCVV 59
DB 67 LNHFNLYNLVTHAIIMVFIMFIMPALFSGFGNYFVPLLGAPDMAPPLNNISEWLLI 126
QY 60 PA-----EGLPGSSRTLY-----HYGVKDLATVEFFYMLVAIIHATIQEYVL 101
DB 127 PAFLLISSTIDGGPGTWTLYPPLNLNCHTGAADVAF-----SLHITGLSSIL 179
QY 102 DKLSRRQL-----TKGKONKLNAGQLSVFVIVSGIWMGMIILASENCLSDPTLLWKSQPH 157
DB 180 GSINLIVTFNNRTPG-----MGLKMPLEFVNSILVTAFLIILAMPVL---SGAI 226
QY 158 NNM---TFQMKFF-----YISQLAYWPHSPPELYF-----QKV---RK-- 189
DB 227 TMLITDRNFGFTTFKPDGGDPLLFQHLFWFPGHPEYIIVILPGFVIGVSVQVSTFSPKPI 286
QY 190 ---QDIPQLIYIGL-----HLFHIGGAY---LLYLNHLGLLLMLHYAVELLSSVCS 236
DB 287 FGXOGMVGAMVIGFVGIWHAHMFVGLSYNALIVFT-AGTMIAVPTGKIFSWIA- 344
QY 237 LLYFGDERYQKGLSLW-----PIVFISGRVLTVLSVVTVGLHLAQT 278
DB 345 -----TWGSGITFPTPMLFAIGFIILEFTIGGVT-GIILSNS 380
RESULT 12
YDIK_ECOLI STANDARD; PRT; 370 AA.
AC P77175;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein ydiK.
GN YDIK OR B1688.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner F.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97251357; PubMed=9097039;
RA Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T.,
RA Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,
RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA Nakade S., Nakamura Y., Nashimoto H., Nishio Y., Oshima T., Saito N.,
RA Sampei G., Seki Y., Sivasubraman S., Tagami H., Takeda J.,
RA Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.;
RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome
corresponding to the 28.0-40.1 min region on the linkage map.";
RL DNA Res. 3:363-377(1996).
CC - SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC - SIMILARITY: BELONGS TO THE UPF0118 (PERM) FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

EMBL; AE000264; AAC74758.1; -;
EMBL; D90811; BAA15450.1; -;
DR EcoGene; EG13970; ydiK.
DR InterPro; IPR002549; UPF0118.
DR Pfam; PF01594; UPF0118; 1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 18 38 POTENTIAL.
FT TRANSMEM 40 60 POTENTIAL.
FT TRANSMEM 66 86 POTENTIAL.
FT TRANSMEM 107 127 POTENTIAL.
FT TRANSMEM 157 177 POTENTIAL.
FT TRANSMEM 214 234 POTENTIAL.
FT TRANSMEM 249 269 POTENTIAL.
FT TRANSMEM 275 295 POTENTIAL.
FT TRANSMEM 309 329 POTENTIAL.
SQ SEQUENCE 370 AA; 39841 MW; 14B7AD79D8C56248 CRC64;
Query Match 5.0%; Score 95; DB 1; Length 370;
Best Local Similarity 19.5%; Pred. No. 0.94;
Matches 67; Conservative 50; Mismatches 86; Indels 140; Gaps 16;
QY 26 MVSCVGMF--FVLGLMFEGTAEMSI--VFLTIQHGVVVPAGLPGSGSRTLHYGVKDLAT 81
DB 24 IVACLWIVQPFILGFAGAGTVVIAWVPLLRLQK-----IMFGRSLAVLVMTLLV 75
QY 82 VFFYMLVAIIHATIQEYVLDKLSRRLQLTGKGNKLNAGQLSVFIVSGIWMGMIILAS 141
DB 76 MVFIIPALLVNS-----IVDGGSLPKAIS 101
QY 142 ENCLSDPTLLWKSQPHNMFTOMKFFIISQLAYWPHSPPELYFQKVRKQDIPGQIYIGL 201

Db 102 SGDMTLPDLA-----WLNTIP-----VIGAKLIAGW 127

Qy 202 H-FHIGGAYLL-----YLN-----HLGLLLMLHYAVELLSSVCSLLYFGDER 244

Db 128 HNLDMGSTALMAKVRPIYIGTTTTFWVGQAAGHIG--RFMVHCALMLIFS--ALLYVRGEQ 183

Qy 245 YOKG-----LSLMPVPIIS-GRLVTLIVSVVTVGLHAGTNRNGNAL 285

Db 184 VAQGIHFATRLAGVRGDAAVLLAAQRAVALGVVVVTVLVAOVLGGI-----GLAV 235

Qy 286 SGNNVLAAXIAVLSSCSIQ-----VYITWLTITVW 317

Db 236 SG-VPYATLTLVLMILSLVQLGCLPLVLPALIIWLTWGTW 277

RESULT 13

SOAL_MOUSE

ID SOAL_MOUSE STANDARD; PRT; 540 AA.

AC Q61263; Q64180;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DE Sterol O-acyltransferase 1 (EC 2.3.1.26) (Cholesterol acyltransferase 1) (Acyl coenzyme A:cholesterol acyltransferase 1) (ACAT-1).

GN SOAT1 OR ACAT.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Liver;

RX MEDLINE=96064687; PubMed=7592824;

RA Uelman P.J., Oka K., Sullivan M.C., Chang T.-Y., Chang C.C.Y.,

RA Chan L.;

RT "Tissue-specific expression and cholesterol regulation of acylcoenzyme

RT A:cholesterol acyltransferase (ACAT) in mice. Molecular cloning of

RT mouse ACAT cDNA, chromosomal localization, and regulation of ACAT in

RT vivo and in vitro.";

RL J. Biol. Chem. 270:26192-26201(1995).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=96158986; PubMed=8579615;

RA Green S., Steinberg D., Quehenberger O.;

RT "Cloning and expression in Xenopus oocytes of a mouse homologue of the

RT human acylcoenzyme A: cholesterol acyltransferase and its potential

RT role in metabolism of oxidized LDL.";

RL Biochem. Biophys. Res. Commun. 218:924-929(1996).

CC -1- FUNCTION: CATALYZES THE FORMATION OF FATTY ACID-CHOLESTEROL

CC ESTERS. PLAYS A ROLE IN LIPOPROTEIN ASSEMBLY AND DIETARY

CC CHOLESTEROL ABSORPTION.

CC -1- CATALYTIC ACTIVITY: Acyl-CoA + cholesterol = CoA + cholesterol

CC ester.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic

CC reticulum.

CC -1- SIMILARITY: BELONGS TO THE STEROL O-ACYLTRANSFERASE FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

CC EMBL; L42293; AAC42075.1; -

CC EMBL; S81092; AAB36050.1; -

CC MGD; MGI:104665; Soatl.

CC InterPro; IPR002688; ACAT.

CC Pfam; PF01800; ACAT; 1.

CC Transferrase; Acyltransferase; Transmembrane; Endoplasmic reticulum;

CC Cholesterol metabolism.

CC TRANSMEM 132 152

FT POTENTIAL.

FT TRANSMEM 311 331 POTENTIAL.

FT TRANSMEM 352 372 POTENTIAL.

FT TRANSMEM 460 480 POTENTIAL.

FT TRANSMEM 488 508 POTENTIAL.

FT CONFLICT 195 195 P > R (IN REF. 2).

SQ SEQUENCE 540 AA; 63739 MW; 8EF900C8BCDF73C0 CRC64;

Query Match 5.0%; Score 94.5; DB 1; Length 540;

Best Local Similarity 21.5%; Pred. No. 1.6;

Matches 45; Conservative 34; Mismatches 93; Indels 37; Gaps 6;

Qy 48 IVFLTLQHGVS-----VVPAGELPSGSRITLYHGVKDLATVFFYMLVALIIHATIQEYV 100

Db 215 LLFLVFLQVLGVFTVYVLLAYTLPPASR-----FILILEQIRLIMKAHSFV 261

Qy 101 LDKLSRRRLQLTKGQNK--LNEAGQLSVFYIVSGIWMIIILASENCLSDPTLLKKSOPHN 158

Db 262 RENIPRVLNAAKEKSSKDPLETVNQYLYF-----LEAPTLLIYRDNYPRTPVWR--GYV 313

Qy 159 MWTQMKEFFYISLAYWPHSPPELYFQKVRKODIPGQLIYIGLHLFHIGGAYLLYLNHLG 218

Db 314 AMQFLQVGCFLFYVYIYIFERLCAPLFRNIKQEPFSARVILCVFNSILPGVLLILFLSFFA 373

Qy 219 LLLMLHYAVELLSSVCSLLYFGDERYQK 247

Db 374 FLHCWLNFAEMLR-----FGDRMFYK 395

RESULT 14

QOXM_SULAC

ID QOXM_SULAC STANDARD; PRT; 788 AA.

AC P39481;

DT 01-FEB-1995 (Rel. 31, Created)

DT 01-FEB-1995 (Rel. 31, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE Quinol oxidase polypeptide I/III (EC 1.9.3.-).

GN SOXM.

OS Sulfolobus acidocaldarius.

OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.

OX NCBI_TaxID=2285;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 33909 / NCIB 11770 / DSM 639;

RX MEDLINE=94357214; PubMed=8076636;

RA Luebben M., Arnaud S., Castresana J., Warne A., Albracht S.P.J.,

RA Saraste M.;

RT "A second terminal oxidase in Sulfolobus acidocaldarius.";

RL Eur. J. Biochem. 224:151-159(1994).

CC -1- FUNCTION: TERMINAL OXIDASE IS THE COMPONENT OF THE RESPIRATORY

CC CHAIN THAT CATALYZES THE REDUCTION OF OXYGEN TO WATER. SOXM

CC FORMS THE FUNCTIONAL CORE OF THE ENZYME COMPLEX.

CC -1- FUNCTION: CATALYTIC SUBUNIT OF THE ENZYME. ELECTRONS ORIGINATING

CC IN A QUINOL ARE TRANSFERRED TO THE BIMETALLIC CENTER OF SOXM

CC FORMED BY A HEME AND COPPER B.

CC -1- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) -> 4 ferrocytochrome

CC c + 2 H(2)O.

CC -1- COFACTOR: TWO HEME GROUPS AND COPPER B.

CC -1- PATHWAY: TERMINAL STEP IN THE RESPIRATORY CHAIN.

CC -1- SUBUNIT: FORMS A COMPLEX WITH AT LEAST SOXC AND A 30 kDa RIESKE

CC FE-S PROTEIN, BUT NEITHER WITH SOXA NOR SOXB.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein.

CC -1- PTM: SOXM IS PROBABLY A PRECURSOR FORM OF SUBUNITS I AND III.

CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE HEME-COPPER

CC RESPIRATORY OXIDASE FAMILY.

CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE CYTOCHROME C

CC OXIDASE SUBUNIT 3 FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

CC EMBL; L42293; AAC42075.1; -

CC EMBL; S81092; AAB36050.1; -

CC MGD; MGI:104665; Soatl.

CC InterPro; IPR002688; ACAT.

CC Pfam; PF01800; ACAT; 1.

CC Transferrase; Acyltransferase; Transmembrane; Endoplasmic reticulum;

CC Cholesterol metabolism.

CC TRANSMEM 132 152

FT POTENTIAL.

entitles requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```
CC EMBL; X73567; CAA51969.1;
CC HSSP; P00396; 10CC.
CC InterPro; IPR000883; COX1.
CC InterPro; IPR000298; CytC_oxdse_III.
CC Pfam; PF00115; COX1; 1.
CC Pfam; PF00510; COX3; 1.
CC PRINTS; PR01165; CYCOXIDASE1.
CC PROSITE; PS00077; COX1; 1.
CC PROSITE; PS0253; COX3; 1.
CC Oxidoreductase; Heme; Copper; Transmembrane; Respiratory chain.
KW TRANSMEM 19 39
FT TRANSMEM 58 78
FT TRANSMEM 100 119
FT TRANSMEM 152 169
FT TRANSMEM 187 206
FT TRANSMEM 232 252
FT TRANSMEM 276 294
FT TRANSMEM 307 326
FT TRANSMEM 337 356
FT TRANSMEM 369 388
FT TRANSMEM 411 432
FT TRANSMEM 449 468
FT TRANSMEM 526 544
FT TRANSMEM 547 564
FT TRANSMEM 614 634
FT TRANSMEM 657 676
FT TRANSMEM 690 707
FT TRANSMEM 729 749
FT TRANSMEM 766 786
FT METAL 62 62
FT METAL 238 238
FT METAL 242 242
FT METAL 288 288
FT METAL 289 289
FT METAL 374 374
FT METAL 376 376
SQ SEQUENCE 788 AA; 87082 MW; 3939C16CDB8A08AD CRC64;

Query Match 5.0%; Score 94.5; DB 1; Length 788;
Best Local Similarity 19.5%; Pred. No. 2.5;
Matches 73; Conservative 57; Mismatches 98; Indels 147; Gaps 19;

Qy 28 SCVG-MFEVLGL--MFECTAEMSIIV-----FLTQ-----HG-----VVVP-AE 62
Db 17 SDVGQYIVLGIVALLIGSVNAALIRQLSFNNAVYDYDAVTLHGIFMFFVWPLST 76
Qy 63 GLPS-----GSTRLYHYGVKDIATVFFYMLVAIIITHATIQE-----YVLDKLSR 106
Db 77 GFANYLVPRMIGAHDLXWPKINALS---FWMVLPVAVILAALISPLLGAVDLGWYMAPLSV 133
Qy 107 RLQLTGKONKLNAGOLSVPIVSG-----IMGMII 138
Db 134 ETTVNYGLGTNL-----IQAILISGLSSYLTGVNFTTKMKKVPYLMKPLFVWGFFT 188
Qy 139 LASENCLSDPTLLMKSQPHNNMTFMKFFIYISQ-----AYNPFHSF 179
Db 189 TAILMIAMPSL-----TAGLVFAYLERLWGPFPFDSDALGSPVLWQQLFWFEGH 238
Qy 180 PELY-----FQVRKQIDIPGOLYIGLHFLHFGGAYL-----LYLNHLGLLLL 222
Db 239 PEVYITILPAMGLVSELLPKMARREIFG---YTAIALSSIAIAFSLAGVMMHH--MFTA 293
Qy 223 MLHAVELLSSVCSL-----LYGEDRYOKGLSL---WPVIFTSGLVT 263
Db 294 IDNTLVQIVSSATWTAIPSGVKLVNWTATLYGGEIRYKTPPTILLISFIVMFLGGITG 353
Qy 264 LIVSVVTVGLHLAGT 278
Db 354 VFFPLVPIDYALNGT 368
```

```
RESULT 15
ATP6_BUCAP
ID ATP6_BUCAP STANDARD; PRT; 272 AA.
AC O51878;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE 30-MAY-2000 (Rel. 39, Last annotation update)
DE ATP synthase A chain (EC 3.6.3.14) (Protein 6).
GN ATP6.
OS Buchnera aphidicola (subsp. Schizaphis graminum).
OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.
OX NCBI_TaxID=98794;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97361981; PubMed=9216881;
RA Clark M.A., Baumann P.;
RT "The (F1F0) ATP synthase of Buchnera aphidicola (endosymbiont of aphids): genetic analysis of the putative ATP operon.";
RL Curr. Microbiol. 35:84-89(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98184963; PubMed=9516544;
RA Clark M.A., Baumann L., Baumann P.;
RT "Sequence analysis of a 34.7-kb DNA segment from the genome of Buchnera aphidicola (endosymbiont of aphids) containing groEL, dnaA, the atp operon, gldA, and rho.";
RL Curr. Microbiol. 36:158-163(1998).
CC -!- FUNCTION: KEY COMPONENT OF THE PROTON CHANNEL; IT MAY PLAY A DIRECT ROLE IN THE TRANSLOCATION OF PROTONS ACROSS THE MEMBRANE (BY SIMILARITY).
CC -!- SUBUNIT: F-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) - THE CATALYTIC CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE SUBUNITS: ALPHA(3), BETA(3), GAMMA(1), DELTA(1), EPSILON(1). CF(0) HAS THREE MAIN SUBUNITS: A, B AND C (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (potential).
CC -!- SIMILARITY: BELONGS TO THE ATPASE A CHAIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
```

```
Query Match 5.0%; Score 94; DB 1; Length 272;
Best Local Similarity 23.5%; Pred. No. 0.8;
Matches 65; Conservative 45; Mismatches 87; Indels 80; Gaps 16;

Qy 73 HYGVKDIATVFFYMLVAIIITHATIQEYVLDKLSRRRLQTLTKGKONKLNAGOLSVPIVSG 132
Db 38 HFWVLNDSIIISLVLCGCFLLSIF--YTVAK-----KITGVPNGLOASIELIFDIFRSN 90
Qy 133 IWGMIIASENCLSDP-----TLLNKSQPHNNMTFMKFF-YISQLAYVHFHSPPELYFOKV 187
Db 91 VKSM--YQGNPLIAPLSLTVFVWVFLNMLDLIPIDFFPFISE--RFFH-EPAM----- 140
```

```

Qy 188 RKQIPGOLIYIGLHFGHAYLLYNHGLGLLLMLHYAVEL--LSSVCSLLYFGDERY 245
Db 141 --RIVPSADINITLSM-----SLGVFILILFYSVKMKGLIGFC----- 176
Qy 246 OKGLSL---WPIVFISG---RLVTLIVSVTVVCLHLAGTNRNCGNALSGN----- 288
Db 177 -KELTLOPFNHPVFFIFNLELVLSLSPISGLRLF-----GNMYSGEMIFILIAGLL 230
Qy 289 -----VNLAAKIAVLSSSCSIQVYITWTLTIVWL 318
Db 231 PWSQFFLNVPWAFHIL--IISIQAFIEMVLTIVYL 265

```

Search completed: September 20, 2002, 14:39:18
Job time: 245 sec

Result No.	Query		Length	DB	ID	Description
	Match	%				
1	1125.5	59.5	374	11	Q91V04	Q91V04 mus musculus
2	1102	58.2	373	13	Q90ZM0	Q90ZM0 xenopus lae
3	1024	54.1	369	13	Q90ZM1	Q90ZM1 brachydanio
4	762	40.3	371	13	Q90ZL9	Q90ZL9 xenopus lae
5	727	38.4	370	11	Q924Z5	Q924Z5 mus musculus
6	566	29.9	368	5	Q9W5C2	Q9W5C2 drosophila
7	566	29.9	368	5	Q9W5C3	Q9W5C3 drosophila
8	561	29.7	368	5	Q9U1L3	Q9U1L3 drosophila
9	441.5	23.3	371	5	Q90J35	Q90J35 caenorhabdi
10	434.5	23.0	373	5	Q9XXV7	Q9XXV7 caenorhabdi
11	348	18.4	159	11	Q9CVJ6	Q9CVJ6 mus musculus
12	163.5	8.6	393	11	Q9D6J1	Q9D6J1 mus musculus
13	138.5	7.3	394	4	Q9HA82	Q9HA82 homo sapien
14	131	6.9	308	10	Q9MG64	Q9MG64 lycopersico
15	119.5	6.3	400	5	Q95RN6	Q95RN6 drosophila
16	116.5	6.2	696	16	Q9CHC2	Q9CHC2 lactococcus


```
Query Match 40.3%; Score 762; DB 13; Length 371;
Best Local Similarity 41.2%; Pred. No. 2.5e-60;
Matches 159; Conservative 67; Mismatches 122; Indels 38; Gaps 7;

QY 1 MGLRKKARPPVLSHEFMVQNHADWVSCVGMFFVL-----GLMFEGETAEMSVFLTLQHG 56
DQ 1 MAFRRR--KSYPLFSQBEFVHNHAD-----IGFFVLCVLGLMFEVTAFTAFILPQYN 54
QY 57 VVPAEGLPSGRRLYHYGVKDLATVFYMLVAIIHATIQEYVLDKLSRRLQTLTKGON 116
DQ 55 SSIGT---LGEILYYHGVKDLVTLFYVVAIIHAIHVOEYILDKLRLSLKVKOS 111
QY 117 KLNAGQLSVFYIVSGTWGMIIASENCISDPTLLKWSQPHNMFTQMFFYISQLAYWF 176
DQ 112 RFNESGQLAAFLHASMFWCLVSVATEGYLSYPTKLWESYHYVLPFOVKFFVLCQLAYWL 171
QY 177 HSPFELYFQVKRQIDPGLIYIGLHFLHIGGAVLLYLNHLGLLMLHYAVELLSSVCS 236
DQ 172 HALPELTFQVKKEEVPRLQYIVLYLLHATAGAYLLNLRGLLILLLOSVAEFLFHAR 231
QY 237 LLXFEDRYQKGLSLWPVIFSGRLVTLIVSVTVVGLHLAGTN-RNGNALSGNVNVIAAK 295
DQ 232 LFYFTDENNORLFWAGVWFVITRLFTLSVLTIGFLARAEBVHTFDPKGTLLNLLFR 291
QY 296 IAVLSSCSIOVYITWTTLTWLQRLWLEDANLHVCGKRRSRKGT-----342
DQ 292 MVVLLMCVSQTMWRFHIFQRLWRE-----CCKEQAAKRSVAVAMKQQAQVVKR 345
QY 343 -----ENGVENPNRIDSPKKKEKAP 363
DQ 346 ESGYHGVVYKAENGSPFQKKIKSP 371

RESULT 5
ID Q92425 ,PRELIMINARY; PRT; 370 AA.
AC Q92425;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE TRAM2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Hartmann E.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY029530; AAK40298.1; -.
SQ SEQUENCE 370 AA; 43182 MW; F8E768AFB0582548 CRC64;

Query Match 38.4%; Score 727; DB 11; Length 370;
Best Local Similarity 40.7%; Pred. No. 3.4e-57;
Matches 151; Conservative 70; Mismatches 134; Indels 16; Gaps 5;

QY 5 KKNARNPPVLSHEFMVQNHADWVSCVGMFFVLGLMFEGETAEMSVFLTLQHG VVPAEGL 64
DQ 4 RRRTKSPLESQBEFVHNHADIGFCLVLCVLGLMFEVTAFTAFILPQYNISVPT--- 60
QY 65 PGSRTLL-YHYGVKDLATVFYMLVAIIHATIQEYVLDKLSRRLQTLTKGONKLEAQ 123
DQ 61 -ADSETVHYHYGPKDLTILFYVITITFIHVAVVOEYILDKISKRLSLKVKHKNESQ 119
QY 124 LSVFYIVSGTWGMIIASENCISDPTLLKWSQPHNMFTQMFFYISQLAYWFHSPFELY 183
DQ 120 LLVFLHSVAVWCFFVIVITGELTNPRSLMEDYPHYVLSYFQVKFFYLGQLAYWLHSLPELY 179
QY 184 FQVKRKQIDPGLIYIGLHFLHIGGAVLLYLNHLGLLMLHYAVELLSSVCSLLYFGDE 243
DQ 184 FQVKRKQIDPGLIYIGLHFLHIGGAVLLYLNHLGLLMLHYAVELLSSVCSLLYFGDE 243
```

```
SQ SEQUENCE 368 AA; 41780 MW; 693794394C2ED787 CRC64;

Query Match
Best Local Similarity 29.9%; Score 566; DB 5; Length 368;
Matches 136; Conservative 64; Mismatches 130; Indels 38; Gaps 9;

QY 4 RKNARNPPVLSHEFPWQNHADWVSCVGMFFVGLMFGTAEISIVFLTQHG--VVPA 61
DB 9 RKTSNKNPILSHEFVIONHADIISCVAMFVVGMLNNESTAAFAFISLHHNVSGEDPS 68
QY 62 EGLPGSRTLYHYGVKDLATVFFYMLVAIIHATQIEVLDKLSRRQLTKGKONKLEA 121
DB 69 RQPYGKPYTYIAGIKDYCAIFFYTLTCIIMHAIIOEFVLDKISKLHLSKFLARFNE 128
QY 122 GOLSVFYIVSGIWMIIASENCISDPTLLWKSQPHNMFTOMKFFYISQLAYWFHSFPE 181
DB 129 GOLVAFYLLSFWGHAHLLKEGYLQVAGLWEGFDPHPSFLHFFYVQLAYLHMLPE 188
QY 182 LYFOKVR-KQDIPGQLIYIGLHLFHIGG-----AYLLYLNHLGLLMLHYAVELLSV 234
DB 189 LYFOKIKKEEQPKIVH-----SISGFTLIVLAYTLISFORLALVLTLYHFSSELLSHV 242
QY 235 CSL--YFGDERYQKGLSLWPIVTSIGRLVTVSVVTVGLHLAGTNRNGNALSNGVNV 292
DB 243 FOLIGVDFREERLAKLRVNNNAVFLIRFATSVIGLTVYIGGV-RSLLAGGLI--- 298
QY 293 AAKI AVLSSCSIQYIYITWTLTVWLQRLWLEDANLHVCGRKRRSRKGTENGVENPNRI 352
DB 299 -----ALQGYLVFSFITEQLRAKRAK-----EAKREAKLALQTKKPAKTPK-- 341
QY 353 DSPPKKKE 360
DB 342 DKVKRKE 349

RESULT 7
Q9W5C3 ID Q9W5C3 PRELIMINARY; PRT; 1575 AA.
AC Q9W5C3
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2001 (TREMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE CG18830 PROTEIN.
GN CG18830 OR CG16994.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Branton R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fodor C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glöck A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris K.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
Liu X., Maitel B., McIntosh T.C., McLeod M.P., McPherson D.,
Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
Palazzo M., Pittman G.S., Pan S., Pollard J.J., Puri V., Reese M.G.,
Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
Zheng X.-H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL; AE003419; AAG23265.1; -.
DR FlyBase; FBgn0040340; EG:BACR7A4.5.
KW Hypothetical protein.
SQ SEQUENCE 1575 AA; 172376 MW; 2F8C0E528B67CD69 CRC64;

Query Match
Best Local Similarity 29.9%; Score 566; DB 5; Length 1575;
Matches 136; Conservative 64; Mismatches 130; Indels 38; Gaps 9;

QY 4 RKNARNPPVLSHEFPWQNHADWVSCVGMFFVGLMFGTAEISIVFLTQHG--VVPA 61
DB 1216 RKTSNKNPILSHEFVIONHADIISCVAMFVVGMLNNESTAAFAFISLHHNVSGEDPS 1275
QY 62 EGLPGSRTLYHYGVKDLATVFFYMLVAIIHATQIEVLDKLSRRQLTKGKONKLEA 121
DB 1276 RQPYGKPYTYIAGIKDYCAIFFYTLTCIIMHAIIOEFVLDKISKLHLSKFLARFNE 1335
QY 122 GOLSVFYIVSGIWMIIASENCISDPTLLWKSQPHNMFTOMKFFYISQLAYWFHSFPE 181
DB 1336 GOLVAFYLLSFWGHAHLLKEGYLQVAGLWEGFDPHPSFLHFFYVQLAYLHMLPE 1395
QY 182 LYFOKVR-KQDIPGQLIYIGLHLFHIGG-----AYLLYLNHLGLLMLHYAVELLSV 234
DB 1396 LYFOKIKKEEQPKIVH-----SISGFTLIVLAYTLISFORLALVLTLYHFSSELLSHV 1449
QY 235 CSL--YFGDERYQKGLSLWPIVTSIGRLVTVSVVTVGLHLAGTNRNGNALSNGVNV 292
DB 1450 FOLIGVDFREERLAKLRVNNNAVFLIRFATSVIGLTVYIGGV-RSLLAGGLI--- 1505
QY 293 AAKI AVLSSCSIQYIYITWTLTVWLQRLWLEDANLHVCGRKRRSRKGTENGVENPNRI 352
DB 1506 -----ALQGYLVFSFITEQLRAKRAK-----EAKREAKLALQTKKPAKTPK-- 1548
QY 353 DSPPKKKE 360
DB 1549 DKVKRKE 1556

RESULT 8
Q9UL13 ID Q9UL13 PRELIMINARY; PRT; 368 AA.
AC Q9UL13
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE EG:BACR7A4.5 PROTEIN.
GN EG:BACR7A4.5 OR CG11642.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
```



```

Db 67 AVEQOEVEHGYLSGILDPAIFFYSCWIVHVAVQYGLDRISKTHLSKYSTFKFG 126
QY 120 EAGQSLSVFYVSGINGMILASENCLSD-----PTLLWKSOP--HNMTFMKFFYISQL 172
Db 127 ESFH-QMEFTVYSIAHAFYIYSER-LEDFSEVKSFSVWLGYPTHERVMSAAYKLIFYFOI 184
QY 173 AYWHSFPFLFYQVKRKIDQPOLYIYGLHFHIGGAYLLYLNHLGLMLLMHYAVELL 232
Db 185 SYWIHQFFEFYLOKLKDEIRQKSVQALHIAFISIAFFNFTRVGLALITLEYITOLIF 244
QY 233 SVCSLLYGGDERYOKGLS-----LWPIVFISGRVTLIVSVVT--VGLHLAGTNRNAGAL 285
Db 245 HIARFAHVG---RKGLSDPAFKFNGSFVLVRLGSIIVIAVTFWYGLRQA-ESPFFVDIS 300
QY 286 SGNVNVLAAKIAVSSSCSIQYITWTTLTVMLQRLWLEDANLHVCGRRKRSRSGTENG 345
Db 301 AGNFNTAVIRNLVNLVAVLLOLELLYSFV-----FHM-GRFESNAKKEKKKS 348
QY 346 VENPNRIDSPPKKKEK 361
Db 349 AAAAAA---PKKEKK 361

RESULT 11
Q9CVJ6 PRELIMINARY; PRT; 159 AA.
AC Q9CVJ6;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DE 1810049E02RIK PROTEIN (FRAGMENT).
GN 1810049E02RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=PANCREAS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikolaio I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Brownstein M.J., Boffelli D., Bojunga N., Carninci P., de Bonaudo M.F.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC Hartmann E.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK013554; BAB28903.1; -
DR EMBL; BC003946; AAH03946.1; -
DR EMBL; AY029531; AAK40299.1; -
DR MGD; MGI:1914510; 2900019C14Rik.
DR InterPro; IPR001356; Homeobox.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS50071; HOMEBOX_2; 1.
SQ SEQUENCE 159 AA; 18053 MW; D88C0B3126B0085A CRC64;

```

```

Query Match 18.4%; Score 348; DB 11; Length 159;
Best Local Similarity 48.18; Pred. No. 1.4e-23;
Matches 76; Conservative 22; Mismatches 50; Indels 10; Gaps 3;
QY 215 NHLGLLLMLHYAVELLSSVGLYFGDERYOKGLSLWPIVFISGRVTLIVSVVTYGLH 274
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

Db 1 NHLGLVLLVLYHVFELFHISRLFYFSDEKYQKGFSLNVLFLVGLRLTLILSVTVQFG 60
QY 275 LA-GNRRNGNALSGNVNVLAAKIAVSSSCSIQYITWTTLTVMLQRLWLEDANLHVCGRRK 333
|| : : : : ||| : : : : ||| : : : : ||| : : : : ||| : : : :
Db 61 LARAENQKLDFTGFNFVNLAVRIAVLASICITQAFMMKFINFQLRRRHSAFQAPVK 120
QY 334 RR-----SRSKKTENGVE---NPNRIDSPPKKKEKA 362
| : : : : ||| : : : : ||| : : : :
Db 121 RRPVATKGRSSRKGTENGVTNGADSPRNRKES 158

RESULT 12
Q9D6J1 PRELIMINARY; PRT; 393 AA.
AC Q9D6J1;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DE 2900019C14RIK PROTEIN (RIKEN CDNA 2900019C14 GENE) (TRH1).
GN 2900019C14RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=HIPPOCAMPUS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikolaio I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaudo M.F.,
RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC Hartmann E.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK013554; BAB28903.1; -
DR EMBL; BC003946; AAH03946.1; -
DR EMBL; AY029531; AAK40299.1; -
DR MGD; MGI:1914510; 2900019C14Rik.
DR InterPro; IPR001356; Homeobox.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS50071; HOMEBOX_2; 1.
SQ SEQUENCE 393 AA; 46016 MW; B2038F7A128F816F CRC64;

```

```

Query Match 8.6%; Score 163.5; DB 11; Length 393;
Best Local Similarity 26.0%; Pred. No. 1.5e-06;
Matches 46; Conservative 40; Mismatches 76; Indels 15; Gaps 5;
QY 106 RLQLTGKQNKLNAGOLSVFYVSGIWMIIASENCLSDPTLLWKSQPHNMATFQMK 165
|| : : : : ||| : : : : ||| : : : : ||| : : : : ||| : : : :
Db 121 RRNRQPSLSKFKCEACWRFVFLQSFVGGTSILYHESLWSPSLCWENYHQTLNLSLY 180
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

GN ASC.
OS Lycopersicon esculentum (Tomato).
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
ON NCBI_TaxID=4081;
OX (1);
RN SEQUENCE FROM N.A.
RP STRAIN=CV. VFNT CHERRY;
RX MEDLINE=99168767; PubMed=10071209;
RA Mesbah L.A., Kneppers T.J., Takken F.L., Laurent P., Hille J.,
RA Nijkamp H.J.;
RA "Genetic and physical analysis of a YAC contig spanning the fungal
RT disease resistance locus Asc of tomato (Lycopersicon esculentum).";
RL Mol. Gen. Genet. 261:50-57(1999).
RN (2);
RN SEQUENCE FROM N.A.
RP STRAIN=CV. VFNT CHERRY;
RX MEDLINE=20243803; PubMed=10781105;
RA Brandwagt B.F., Mesbah L.A., Takken F.L.W., Laurent P.L.,
RA Kneppers T.J.A., Hille J., Nijkamp H.J.J.;
RA "A longevity assurance gene homolog of tomato mediates resistance to
RT Alternaria alternata f. sp. lycopersici toxins and fumonisin B1.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:4961-4966(2000).
DR EMBL; AF198177; AAF67518.1;
SQ SEQUENCE 308 AA; 36301 MW; 2BD3378CE53F416D CRC64;

Query Match 6.9%; Score 131; DB 10; Length 308;
Best Local Similarity 19.5%; Pred. No. 0.00094;
Matches 46; Conservative 50; Mismatches 80; Indels 60; Gaps

QY 62 EGPLSGSRTLYHYGVKDLATVPFYMLVAIIIIHATIQEYVLDKLSRRLQTK----- 112
DB 16 ESLPE-----YQDLIFLFFALFFPVILRILDRFVEALAKEMIFGKKTVVNNGR 66

QY 113 ---GKQNKLINEAGQLSVFYIVSGIWGMILASENCUSDPTLLWKSQ-----PHNMWTFQM 164
DB 67 EERKTKINKFKESAWKVFYLSAELLALSVCNEPWTDSRYFWAGPGDVVWPLKMKKL 126

QY 165 KFYVISOLAWFHS-FPELYFQVRKODIPGOLYIYGLHLFHIGGAYLLVLNH-----L 217
DB 127 KLYXWVAGGFYFSIATLYWE--TRYDPAQIIH-----HVTYVSLVLSVYVGFARI 179

QY 218 GLLLMLHYAVELLSSVCSL-----LYFGDERYQKGLSWPIVFIGSLRV 262
DB 180 GSVVALHDGSDVFMETIAKMSKYSFGDLIADIFF-----SLFALVETSLRII 226

RESULT 15
ID Q95RN6 PRELIMINARY; PRT: 400 AA.
AC Q95RN6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE LD18904P.
OS CG13898.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Orthopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidae; Drosophilidae; Drosophila.
ON NCBI_Taxid=7227;
RN (1);
RN SEQUENCE FROM N.A.
RP STRAIN=Y. CN BW SP.
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Erise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RA Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY061255; AAL28803.1;

```

Search completed: September 20, 2002, 14:38:59
Job time: 246 sec